

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 15:39:22 ; Search time 13.125 Seconds
(without alignments)
131.920 Million cell updates/sec

Title: US-10-005-684-7
Perfect score: 92
Sequence: 1 LSIALHVGDFHFEQLLSG 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 92 | 100.0 | 558 | 1 NUNMS | glucose-6-phosphat |
| 2 | 85 | 92.4 | 553 | 2 I47142 | glucose-6-phosphat |
| 3 | 85 | 92.4 | 558 | 1 NUPG | glucose-6-phosphat |
| 4 | 84 | 91.3 | 558 | 2 I48073 | glucose-6-phosphat |
| 5 | 79 | 85.9 | 548 | 2 AF0452 | glucose-6-phosphat |
| 6 | 70 | 76.1 | 613 | 2 T27469 | hypothetical prote |
| 7 | 63 | 68.5 | 557 | 1 NUVKL | glucose-6-phosphat |
| 8 | 62 | 67.4 | 563 | 2 F64130 | glucose-6-phosphat |
| 9 | 61 | 66.3 | 549 | 1 NUBC | glucose-6-phosphat |
| 10 | 61 | 66.3 | 549 | 2 H91254 | glucose-6-phosphat |
| 11 | 61 | 66.3 | 549 | 2 D86095 | glucose-6-phosphat |
| 12 | 61 | 66.3 | 549 | 2 AD1013 | glucose-6-phosphat |
| 13 | 60 | 65.2 | 550 | 2 T43196 | glucose-6-phosphat |
| 14 | 60 | 65.2 | 550 | 2 T39509 | probable glucose-6 |
| 15 | 60 | 65.2 | 554 | 1 NUBY | glucose-6-phosphat |
| 16 | 58 | 63.0 | 549 | 2 G84996 | glucose-6-phosphat |
| 17 | 57 | 62.0 | 562 | 2 T48970 | glucose-6-phosphat |
| 18 | 56 | 60.9 | 557 | 2 T48443 | glucose-6-phosphat |
| 19 | 54 | 58.7 | 550 | 2 B82330 | glucose-6-phosphat |
| 20 | 51 | 55.4 | 584 | 2 T14631 | glucose-6-phosphat |
| 21 | 49 | 53.3 | 554 | 2 H83053 | glucose-6-phosphat |
| 22 | 47 | 51.1 | 353 | 2 B72738 | glucose-6-phosphat |
| 23 | 45 | 48.9 | 502 | 2 H82832 | hypothetical prote |
| 24 | 45 | 48.9 | 541 | 2 T00031 | glucose-6-phosphat |
| 25 | 45 | 48.9 | 744 | 2 T75358 | glucose-6-phosphat |
| 26 | 44 | 47.8 | 396 | 2 B70394 | oxetanocin A biosy |
| 27 | 44 | 47.8 | 605 | 2 F84109 | ABC transporter - |
| 28 | 43 | 46.7 | 430 | 1 HMIHMS | two-component sens |
| 29 | 43 | 46.7 | 439 | 2 JQ0997 | hemagglutinin-este |

| | | | | | |
|----|----|------|------|----------|--------------------|
| 30 | 43 | 46.7 | 607 | 1 NUUTB | glucose-6-phosphat |
| 31 | 43 | 46.7 | 723 | 2 A84425 | hypothetical prote |
| 32 | 43 | 46.7 | 735 | 2 T30096 | hypothetical prote |
| 33 | 42 | 45.7 | 283 | 2 E70038 | maltoextrin trans |
| 34 | 42 | 45.7 | 355 | 1 GBBPP4 | beta protein - sat |
| 35 | 42 | 45.7 | 427 | 1 HMIHMH | hemagglutinin-este |
| 36 | 42 | 45.7 | 511 | 2 C95205 | hypothetical prote |
| 37 | 42 | 45.7 | 818 | 2 S76976 | pyruvate, water di |
| 38 | 42 | 45.7 | 1167 | 2 T13927 | adenylate cyclase |
| 39 | 41 | 44.6 | 225 | 2 T36587 | hypothetical prote |
| 40 | 41 | 44.6 | 240 | 2 AB0499 | probable phosphat |
| 41 | 41 | 44.6 | 507 | 2 A39411 | glucose-6-phosphat |
| 42 | 41 | 44.6 | 537 | 2 A38170 | cytochrome d compl |
| 43 | 41 | 44.6 | 539 | 2 E87276 | glucose-6-phosphat |
| 44 | 41 | 44.6 | 589 | 2 S42385 | hypothetical prote |
| 45 | 41 | 44.6 | 636 | 2 S70877 | hypothetical prote |

ALIGNMENTS

RESULT 1

NUNMS

Glucose-6-phosphate isomerase (EC 5.3.1.9) - mouse
N;Alternate names: neuroleukin; phosphoglucose isomerase; phosphohexose isomerase
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 24-Nov-1999
C;Accession: A24439; I49701
R;Gurney, M.E.; Heinrich, S.P.; Lee, M.R.; Yin, H.
Science 234, 566-574, 1986
A;Title: Molecular cloning and expression of neuroleukin, a neurotrophic factor for spin
A;Reference number: A24439; MUID:87018838; PMID:3764429
A;Accession: A24439
A;Molecule type: mRNA
A;Residues: 1-558 <RES>
A;Cross-references: GB:M14220; NID:G200064; PIDN:AAA39825.1; PID:G200065
A;Experimental source: salivary
R;Faik, P.; Walker, J.L.; Morgan, M.J.
Genomics 21, 122-127, 1994
A;Title: Identification of a novel tandemly repeated sequence present in an intron of th
A;Reference number: I49701; MUID:94375004; PMID:7545951
A;Accession: I49701
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 269-558 <RES>
A;Cross-references: GB:I00104; NID:G309267; PIDN:AAA65641.1; PID:G309268
C;Comment: This protein, found in a wide variety of tissues, is a potent, pleiotrophic g
sympathetic or parasympathetic neurons.
C;Comment: As a lymphokine produced by lectin-stimulated lymphocytes, neuroleukin stimu
C;Superfamily: glucose-6-phosphate isomerase
C;Keywords: blocked amino end; gluconeogenesis; glycolysis; growth factor; homodimer; in
F;2-558/Product: glucose-6-phosphate isomerase #status predicted <MAR>
F;2/Modified site: blocked amino end (Ala) (in mature form) #status experimental
F;519/Active site: Lys #status predicted

Query Match 100.0%; Score 92; DB 1; Length 558;
Best Local Similarity 100.0%; Pred.No. 3.9e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHFEQLLSG 18
|||||
DB 282 LSIALHVGDFHFEQLLSG 299

RESULT 2

I47142

Glucose-6-phosphate isomerase (EC 5.3.1.9) - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jul-2000
C;Accession: I47142
R;Claes, V.; Kettmann, R.; Burny, A.
Gene 150, 235-241, 1994
A;Title: Structure of the gene encoding pig phosphoglucose isomerase.

```

A:Reference number: I47142; MUID:95121909; PMID:7821788
A:Accession: I47142
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-553 <CLA>
A:Cross-references: EMBL:Z28396; NID:9433921; PIDN:CAA82246.1; PID:g1364248
C:Genetics:
A:Introns: 41/2, 71/3, 94/3, 134/3, 160/3, 209/3, 233/3, 248/3, 266/3, 287/1, 301/3, 352
C:Superfamily: glucose-6-phosphate isomerase
C:Keywords: intramolecular oxidoreductase; isomerase

Query Match          92.4%; Score 85; DB 2; Length 553;
Best Local Similarity 94.4%; Pred. No. 5.4e-06;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSIALHVGDFHFEQLLSG 18
|||||:|||||
Db 280 LSIALHVGDFNFEQLLSG 297

RESULT 3
NUPG
Glucose-6-phosphate isomerase (EC 5.3.1.9) - pig
N:Alternate names: neuroleukin; phosphoglucose isomerase; phosphohexose isomerase
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 02-Jun-2000
C:Accession: S00895; V13506; I46587
R:Chaput, M.; Claes, V.; Portetelle, D.; Cludts, I.; Cravador, A.; Burny, A.; Gras, H.;
Nature 332, 454-455, 1988
A:Title: The neurotrophic factor neuroleukin is 90% homologous with phosphohexose isomerase
A:Reference number: S00895; MUID:88175070; PMID:332744
A:Accession: S00895
A:Molecule type: mRNA
A:Residues: 1-558 <CHA>
A:Cross-references: EMBL:X07382; NID:G2042; PIDN:CAA30295.1; PID:g2043
A:Note: part of this sequence was confirmed by protein sequencing
R:Claes, V.; Taquet, A.N.; Kettmann, R.; Burny, A.
Biochim. Biophys. Acta 1087, 339-340, 1990
A:Title: Sequence analysis of the pig phosphoglucose isomerase gene promoter region.
A:Reference number: S13506; MUID:91064390; PMID:2248981
A:Accession: S13506
A:Molecule type: mRNA
A:Residues: 1-17 <CLA>
A:Cross-references: EMBL:X53719; NID:g1848; PIDN:CAA37755.1; PID:g1849
R:Davies, W.; Harbitz, J.; Hauge, J.G.
Anim. Genet. 18, 233-240, 1987
A:Title: A partial cDNA clone for porcine glucosephosphate isomerase: Isolation, character
A:Reference number: I46587; MUID:88021976; PMID:2889410
A:Accession: I46587
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 238-343, 'H', 345-357 <DAV>
A:Cross-references: GB:M54975; NID:g164481; PIDN:AAA31048.1; PID:g552364
C:Comment: Neuroleukin promotes neuritic sprouting from motor neurons but not from sympa
C:Comment: As a lymphokine produced by lectin-stimulated lymphocytes, neuroleukin stimu
C:Genetics:
A:Gene: PGI; gpi
C:Superfamily: glucose-6-phosphate isomerase
C:Keywords: blocked amino end; gluconeogenesis; glycolysis; growth factor; homodimer; in
F12-558/Product: glucose-6-phosphate isomerase #status predicted <MAN>
F12/Modified site: blocked amino end (ala) (in mature form) #status experimental
F1519/Active site: Lys #status predicted

Query Match          92.4%; Score 85; DB 1; Length 558;
Best Local Similarity 94.4%; Pred. No. 5.4e-06;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSIALHVGDFHFEQLLSG 18
|||||:|||||
Db 282 LSIALHVGDFNFEQLLSG 299

RESULT 4
A:Reference number: I47142; MUID:95121909; PMID:7821788
A:Accession: I47142
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-553 <CLA>
A:Cross-references: EMBL:Z28396; NID:9433921; PIDN:CAA82246.1; PID:g1364248
C:Genetics:
A:Introns: 41/2, 71/3, 94/3, 134/3, 160/3, 209/3, 233/3, 248/3, 266/3, 287/1, 301/3, 352
C:Superfamily: glucose-6-phosphate isomerase
C:Keywords: intramolecular oxidoreductase; isomerase

Query Match          91.3%; Score 84; DB 2; Length 558;
Best Local Similarity 88.9%; Pred. No. 7.9e-06;
Matches 16; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSIALHVGDFHFEQLLSG 18
|||||:|||||
Db 282 LSIALHVGDFNFEQLLSG 299

RESULT 5
AF0452
Glucose-6-phosphate isomerase (EC 5.3.1.9) [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C:Accession: AF0452
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; F
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, E
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AF0452
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-548 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC93186.1; PID:g15981635; GSPDB:GN00175
C:Genetics:
A:Gene: Pgi
C:Superfamily: glucose-6-phosphate isomerase
C:Keywords: intramolecular oxidoreductase; isomerase

Query Match          85.9%; Score 79; DB 2; Length 548;
Best Local Similarity 88.9%; Pred. No. 5.1e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LSIALHVGDFHFEQLLSG 18
|||||:|||||
Db 279 LSIALSVGFHFHFEQLLSG 296

RESULT 6
T27469
Hypothetical protein Y87G2A.q - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C:Accession: T27469
R:White, S.
Submitted to the EMBL Data Library, September 1999
A:Reference number: Z20371
A:Accession: T27469
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-613 <WIL>
A:Cross-references: EMBL:AL110500; PIDN:CAB54491.1; CESP:Y87G2A.q
A:Experimental source: clone Y87G2A

```

C:Genetics:

A:Gene: CESP:Y87G2A.g
A:Introns: 40/2; 138/1; 447/1; 454/2; 483/3; 574/2
C:Superfamily: Glucose-6-phosphate isomerase

Query Match 76.1%; Score 70; DB 2; Length 613;
Best Local Similarity 66.7%; Pred. No. 0.0017;
Matches 12; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSIALHVGFDHFEQLLSG 18
|||||:||||:||||
Db 280 LSIATVHIGFDNYEKLDG 297
|||||:||||:||||

RESULT 7

NUVKL

Glucose-6-phosphate isomerase (EC 5.3.1.9) - Yeast (*Kluyveromyces marxianus* var. *lactis*)
N:Alternate names: phosphoglucose isomerase; phosphohexose isomerase
C:Species: *Kluyveromyces marxianus* var. *lactis*, *Candida sphærica*
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 18-Jun-1999
C:Accession: S01414
R:Wesolowski-Louvel, M.; Goffrini, P.; Ferrero, I.
Nucleic Acids Res. 16, 8714, 1988
A:Title: The RAG2 gene of the Yeast *Kluyveromyces lactis* codes for a putative phosphoglu
A:Reference number: S01414; UID:88335624; PMID:3419932
A:Accession: S01414
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-557 <WES>
A:Cross-references: EMBL:X12360; NID:g2892; PIDN:CAA30923.1; PID:g2893
C:Genetics:
A:Gene: RAG2
C:Superfamily: glucose-6-phosphate isomerase
C:Keywords: gluconeogenesis; glycolysis; homodimer; intramolecular oxidoreductase; isom
F:523/Active site: Lys #status predicted

Query Match 68.5%; Score 63; DB 1; Length 557;
Best Local Similarity 61.1%; Pred. No. 0.021;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSIALHVGFDHFEQLLSG 18
|||||:||||:||||
Db 294 LSVLVIGFDNFEPFLKG 311
|||||:||||:||||

RESULT 8

F64130

Glucose-6-phosphate isomerase (EC 5.3.1.9) - *Haemophilus influenzae* (strain Rd KW20)
C:Species: *Haemophilus influenzae*
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 16-Jul-1999
C:Accession: F64130
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.
A:Reference number: A64000; UID:95350630; PMID:7542800
A:Accession: F64130
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-563 <TIGR>
A:Cross-references: GB:U32831; GB:L42023; NID:G1574407; PIDN:AAC23219.1; PID:G1574413; T
C:Genetics:
A:Start codon: GTG
C:Superfamily: Glucose-6-phosphate isomerase
C:Keywords: gluconeogenesis; glycolysis; intramolecular oxidoreductase; isomerase
F:529/Active site: Lys #status predicted

Query Match 67.4%; Score 62; DB 2; Length 563;
Best Local Similarity 66.7%; Pred. No. 0.031;
Matches 12; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSIALHVGFDHFEQLLSG 18
|||||:||||:||||
Db 293 LSIATLSIGFENFALLNG 310
|||||:||||:||||

RESULT 9

NUEC

Glucose-6-phosphate isomerase (EC 5.3.1.9) - *Escherichia coli* (strain K-12)
N:Alternate names: phosphoglucose isomerase; phosphohexose isomerase
C:Species: *Escherichia coli*
C>Date: 31-Mar-1990 #sequence_revision 17-Oct-1997 #text_change 01-Mar-2002
C:Accession: H65209; JS0142; S04396
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A:Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of *Escherichia coli* K-12.
A:Reference number: A64720; UID:97426617; PMID:9278503
A:Accession: H65209
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-549 <BLAT>
A:Cross-references: GB:AE000476; GB:U00096; NID:G1790456; PIDN:AAC76995.1; PID:G1790457;
A:Experimental source: strain K-12, substrain MG1655
R:Froman, B.E.; Tait, R.C.; Gottlieb, L.D.
Mol. Gen. Genet. 217, 126-131, 1989
A:Title: Isolation and characterization of the phosphoglucose isomerase gene from *Escher*
A:Reference number: JS0142; UID:89364675; PMID:2549364
A:Accession: JS0142
A:Molecule type: DNA
A:Residues: 1-316, 'V', 318-549 <PRO>
A:Cross-references: GB:X15196; NID:G42376; PIDN:CAA33268.1; PID:G42377
A:Experimental source: strain JM101
A:Note: the authors translated the codon CAG for residue 8 as Trp
C:Comment: this enzyme catalyzes the reversible isomerization of glucose-6-phosphate and
C:Genetics:
A:Gene: pgi
A:Map position: 91 min
C:Superfamily: glucose-6-phosphate isomerase
C:Keywords: glycolysis; homodimer; intramolecular oxidoreductase; isomerase
F:514/Active site: Lys #status predicted

Query Match 66.3%; Score 61; DB 1; Length 549;
Best Local Similarity 66.7%; Pred. No. 0.045;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSIALHVGFDHFEQLLSG 18
|||||:||||:||||
Db 279 LSVLVIGFDNFVLLSG 296
|||||:||||:||||

RESULT 10

H91254

Glucosephosphate isomerase [imported] - *Escherichia coli* (strain O157:H7, substrain RIMC
C:Species: *Escherichia coli*
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: H91254
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and Genc
A:Reference number: A99629; UID:21156231; PMID:11258796
A:Accession: H91254
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-549 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA038431.1; PID:G13364485; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509552
C:Genetics:
A:Gene: ECs5008
C:Superfamily: glucose-6-phosphate isomerase

Query Match 66.3%; Score 61; DB 2; Length 549;
Best Local Similarity 66.7%; Pred. No. 0.045;

Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHFEQLLSG 18
 ||| :||| :|||
 Db 279 LSIVLSIGDFNFVELLSG 296

RESULT 11
 D86095
 Glucose-6-phosphate isomerase [imported] - Escherichia coli (strain O157:H7, substrain EDL958)
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: D86095
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: D86095
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-549 <STO>
 A:Cross-references: GB:AB005174; NID:gl2518968; PIDN:AGS5224.1; GSPDB:GN00145; UWGP:Z56
 A:Experimental source: strain O157:H7, substrain EDL958
 C:Genetics:
 A:Gene: pgi
 C:Superfamily: glucose-6-phosphate isomerase

Query Match 66.3%; Score 61; DB 2; Length 549;
 Best Local Similarity 66.7%; Pred. No. 0.045;
 Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHFEQLLSG 18
 ||| :||| :|||
 Db 279 LSIVLSIGDFNFVELLSG 296

RESULT 12
 AD1013
 Glucose-6-phosphate isomerase (EC 5.3.1.9) [imported] - Salmonella enterica subsp. enter
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AD1013
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 , S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AD1013
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-549 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD09205.1; PID:gl6505209; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY4417
 C:Superfamily: glucose-6-phosphate isomerase
 C:Keywords: intramolecular oxidoreductase; isomerase

Query Match 66.3%; Score 61; DB 2; Length 549;
 Best Local Similarity 72.2%; Pred. No. 0.045;
 Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHFEQLLSG 18
 ||| :||| :|||
 Db 279 LSIVLSIGDFNFVELLSG 296

RESULT 13
 T43196
 probable glucose-6-phosphate isomerase (EC 5.3.1.9) - fission yeast (Schizosaccharomyces

C:Species: Schizosaccharomyces pombe
 C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
 C:Accession: T43196
 R:Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
 DNA Res. 4, 363-369, 1997
 A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
 A:Reference number: Z17323; MUID:98162722; PMID:9501991
 A:Accession: T43196
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-550 <YOS>
 A:Cross-references: EMBL:D89268; NID:gl749743; PIDN:BAA1929.1; PID:gl749744
 A:Experimental source: strain PR745
 C:Function:
 A:Description: catalyzes the reversible isomerization of glucose-6-phosphate and fructose
 A:Pathway: glycolysis; gluconeogenesis
 C:Superfamily: glucose-6-phosphate isomerase
 C:Keywords: gluconeogenesis; glycolysis; intramolecular oxidoreductase; isomerase

Query Match 65.2%; Score 60; DB 2; Length 550;
 Best Local Similarity 55.6%; Pred. No. 0.065;
 Matches 10; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHFEQLLSG 18
 ||| :||| :|||
 Db 287 LSVALYIGYENFEFLSG 304

RESULT 14
 T39509
 glucose-6-phosphate isomerase, cytosolic - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T39509
 R:Beck, A.; Reinhardt, R.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, December 1998
 A:Reference number: Z21859
 A:Accession: T39509
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-550 <BEC>
 A:Cross-references: EMBL:AL034433; PIDN:CAA22338.1; GSPDB:GN00066; SPDB:SPBCL604.05
 A:Experimental source: strain 972h; cosmid c1604
 C:Genetics:
 A:Gene: SPDB:SPBCL604.05
 A:Map position: 1
 C:Superfamily: glucose-6-phosphate isomerase

Query Match 65.2%; Score 60; DB 2; Length 550;
 Best Local Similarity 55.6%; Pred. No. 0.065;
 Matches 10; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHFEQLLSG 18
 ||| :||| :|||
 Db 287 LSVALYIGYENFEFLSG 304

RESULT 15
 NUBY
 Glucose-6-phosphate isomerase (EC 5.3.1.9) - Yeast (Saccharomyces cerevisiae)
 N:Alternate names: phosphoglucose isomerase; phosphohexose isomerase; protein YBR1407; p
 C:Species: Saccharomyces cerevisiae
 C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 21-Jul-2000
 C:Accession: J70484; A40073; S34021; S45445; S46068
 R:Tekamp-Olson, P.; Najarian, R.; Burke, R.L.
 Gene 73, 153-161, 1988
 A:Title: The isolation, characterization and nucleotide sequence of the phosphoglucose
 A:Reference number: J70484; MUID:89211945; PMID:3072254
 A:Accession: J70484
 A:Molecule type: DNA
 A:Residues: 1-554 <TEK>
 A:Cross-references: GB:M21696; NID:gl72224; PIDN:AAA34894.1; PID:gl72225
 R:Green, J.B.A.; Wright, A.P.H.; Cheung, W.Y.; Lancashire, W.E.; Hartley, B.S.

Mol. Gen. Genet. 215, 100-106, 1988
 A;Title: The structure and regulation of phosphoglucose isomerase in *Saccharomyces cerevisiae*
 A;Reference number: A40073; MUID:89201230; PMID:3071735
 A;Accession: A40073
 A;Molecule type: DNA
 A;Residues: 1-554 <GRE>
 A;Cross-references: GB:M37267; NID:g172141; PIDN:AAA34862.1; PID:g172142
 R;Demolis, N.; Mallet, L.; Bussereau, F.; Jacquet, M.
 Yeast 9, 645-659, 1993
 A;Title: RIM2, MGI1 and PG11 are located within an 8 kb segment of *Saccharomyces cerevisiae* leucine zipper motif.
 A;Reference number: S33966; MUID:93348777; PMID:8346681
 A;Accession: S34021
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-554 <DEM>
 A;Cross-references: EMBL:Z21487; NID:g311665; PIDN:CAA79683.1; PID:g311672
 R;Mallet, L.; Bussereau, F.; Jacquet, M.
 Yeast 10, 819-831, 1994
 A;Title: Nucleotide sequence analysis of an 11.7 kb fragment of yeast chromosome II including the *PGI1* gene.
 A;Reference number: S45445; MUID:95066384; PMID:7975899
 A;Accession: S45445
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-554 <VAL>
 A;Cross-references: EMBL:Z21487; NID:g311665; PIDN:CAA79683.1; PID:g311672
 R;Bussereau, F.; Demolis, N.; Jacquet, M.; Mallet, L.
 submitted to the Protein Sequence Database, August 1994
 A;Reference number: S46054
 A;Accession: S46054
 A;Molecule type: DNA
 A;Residues: 1-554 <BUS>
 A;Cross-references: EMBL:Z36065; NID:g536564; PIDN:CAA85158.1; PID:g536565; GSPDB:GN000000
 C;Comment: This enzyme catalyzes the interconversion of glucose-6-phosphate and fructose-6-phosphate.
 C;Genetics:
 A;Gene: SGD:PGI1; MIPS:YBR196C
 A;Cross-references: SGD:S0000400; MIPS:YBR196C
 A;Map position: 2R
 C;Superfamily: glucose-6-phosphate isomerase
 C;Keywords: gluconeogenesis; glycolysis; homodimer; intramolecular oxidoreductase; isomerase
 F:520/Active site: Lys #status predicted

Query Match 65.2%; Score 60; DB 1; Length 554;
 Best Local Similarity 55.8%; Pred. No. 0.066;
 Matches 10; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSTALHVGFDHFEQLLSG 18
 |||::|::|::|
 Db 291 LSVLYIGYDNEAFLLKG 308

Search completed: May 19, 2004, 15:47:12
 Job time : 14.125 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 15:38:37 ; Search time 37.5 Seconds
(without alignments)
151.449 Million cell updates/sec

Title: US-10-005-684-7

Perfect score: 92

Sequence: 1 LSIALHVGDFHFEQLLSG 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriophage:*

17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Match | Query Match | Description |
|------------|-------|---------|-------------|-------------|
| 1 | 92 | 100.0 | 475 | 11 Q8C675 |
| 2 | 88 | 95.7 | 159 | 11 Q89062 |
| 3 | 85 | 92.4 | 85 | 6 Q46595 |
| 4 | 85 | 92.4 | 186 | 6 Q95M65 |
| 5 | 85 | 92.4 | 281 | 6 Q8HY2 |
| 6 | 85 | 92.4 | 300 | 4 Q9BRD3 |
| 7 | 84 | 91.3 | 553 | 13 Q7ZU30 |
| 8 | 80 | 87.0 | 553 | 13 Q8QFT2 |
| 9 | 80 | 87.0 | 553 | 13 Q8QF02 |
| 10 | 78 | 84.8 | 553 | 13 Q8QF06 |
| 11 | 77 | 83.7 | 553 | 13 Q8QF07 |
| 12 | 75 | 81.5 | 553 | 13 Q8QFT1 |
| 13 | 71 | 77.2 | 554 | 5 Q8T313 |
| 14 | 70 | 76.1 | 551 | 5 Q9UIQ2 |
| 15 | 70 | 76.1 | 553 | 13 Q8QF01 |
| 16 | 63 | 68.5 | 562 | 5 Q9GRF4 |

| | | | | | |
|----|----|------|------|----|--------|
| 17 | 62 | 67.4 | 536 | 3 | Q874Q4 |
| 18 | 61 | 66.3 | 326 | 2 | Q9KH84 |
| 19 | 61 | 66.3 | 326 | 2 | Q9KH86 |
| 20 | 61 | 66.3 | 326 | 2 | Q9K2T3 |
| 21 | 61 | 66.3 | 326 | 2 | Q9K2Q3 |
| 22 | 61 | 66.3 | 326 | 2 | Q9K328 |
| 23 | 61 | 66.3 | 326 | 2 | Q9K376 |
| 24 | 61 | 66.3 | 326 | 2 | Q9KH85 |
| 25 | 61 | 66.3 | 326 | 2 | Q9K2R7 |
| 26 | 61 | 66.3 | 326 | 2 | Q9KH87 |
| 27 | 61 | 66.3 | 549 | 16 | Q7UBD1 |
| 28 | 61 | 66.3 | 551 | 16 | Q83IN9 |
| 29 | 59 | 64.1 | 547 | 16 | Q7VNR9 |
| 30 | 58 | 63.0 | 550 | 2 | Q83WT9 |
| 31 | 57 | 62.0 | 508 | 5 | Q8SRY1 |
| 32 | 54 | 58.7 | 544 | 16 | Q82SP4 |
| 33 | 54 | 58.7 | 553 | 2 | Q84CX7 |
| 34 | 53 | 57.6 | 558 | 16 | Q7VR14 |
| 35 | 51 | 55.4 | 346 | 5 | Q86LQ7 |
| 36 | 51 | 55.4 | 346 | 5 | Q86LQ5 |
| 37 | 51 | 55.4 | 346 | 5 | Q867P9 |
| 38 | 51 | 55.4 | 346 | 5 | Q867E0 |
| 39 | 51 | 55.4 | 608 | 5 | Q61113 |
| 40 | 50 | 54.3 | 554 | 2 | Q84814 |
| 41 | 48 | 52.2 | 1230 | 16 | Q828U1 |
| 42 | 47 | 51.1 | 200 | 5 | Q867P4 |
| 43 | 47 | 51.1 | 346 | 5 | Q86LR1 |
| 44 | 47 | 51.1 | 346 | 5 | Q86LR0 |
| 45 | 47 | 51.1 | 346 | 5 | Q86LQ9 |

ALIGNMENTS

RESULT 1

Q8C675 PRELIMINARY; PRT; 475 AA.

AC Q8C675; DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 23, Last annotation update)

DE Glucose phosphate isomerase 1 complex.

GN GPII.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Head;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

60,770 full-length cDNAs."

RL Nature 420:563-573(2002).

DR EMBL; AK076424; BAC3635.1; --

DR PIR; PT0534; PT0534

DR MGD; MGI:95797; GpiI.

DR GO; GO:0004347; F:Glucose-6-phosphate isomerase activity; IEA.

DR GO; GO:0006094; P:Glucosyltransferase; IEA.

DR GO; GO:0006096; P:Glycolysis; IEA.

DR InterPro; IPR001672; G6P_Isomerase.

DR Pfam; PF00342; PGI; 1.

DR PRINTS; PR00662; G6PISOMERASE.

DR PROSITE; PS00765; P_GLUCOSE ISOMERASE 1; 1.

DR PROSITE; PS00174; P_GLUCOSE ISOMERASE 2; 1.

SQ SEQUENCE 475 AA; 53163 MW; D2478D514DB8063B CRC64;

Query Match

Best Local Similarity 100.0%; Score 92; DB 11; Length 475;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHFEQLLSG 18
 |||||
 Db 199 LSIALHVGDFHFEQLLSG 216

RESULT 2
 ID O89062 PRELIMINARY; PRT; 159 AA.
 AC O89062;
 DT 01-NOV-1998 (TREMELrel. 08, Created)
 DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Glucose-6-phosphate isomerase (EC 5.3.1.9) (GPI) (Phosphoglucose
 isomerase) (PGI) (Phosphohexose isomerase) (PHI) (Fragment).
 GN GPI.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Spleen;
 RX MEDLINE=9901297; PubMed=9798653;
 RA Chu C.C., Paul W.E.;
 RT "Expressed genes in interleukin-4 treated B cells identified by cDNA
 representation difference analysis.";
 RL Mol. Immunol. 35:487-502(1998).
 CC -1- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE = D-FRUCTOSE 6-
 PHOSPHATE.
 CC -1- PATHWAY: INVOLVED IN GLYCOLYSIS AND IN GLUCONEOGENESIS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GPI FAMILY.
 DR EMBL; U89408; AAC36515.1; -.
 DR HSP; Q9NIE2; IHOX.
 DR MGD; MGI:95797; Gpi1.
 DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
 DR GO; GO:0016853; F:isomerase activity; IEA.
 DR GO; GO:0006094; P:gluconeogenesis; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR001672; G6P_Isomerase.
 DR Pfam; PF00342; FGI; 1.
 DR PRINTS; PR00662; G6PISOMERASE.
 DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
 KW Gluconeogenesis; Glycolysis; Isomerase.
 FT NON_TER 1 1
 FT NON_TER 159 159
 SQ SEQUENCE 159 AA; 17892 MW; B4F0BA9772043331 CRC64;

Query Match 95.7%; Score 88; DB 11; Length 159;
 Best Local Similarity 94.4%; Pred. No. 1.3e-06;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHFEQLLSG 18
 |||||
 Db 48 LSIVLHVGDFHFEQLLSG 65

RESULT 3
 ID O46595 PRELIMINARY; PRT; 85 AA.
 AC O46595;
 DT 01-JUN-1998 (TREMELrel. 06, Created)
 DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Glucosephosphate isomerase (EC 5.3.1.9) (Glucose-6-phosphate
 isomerase) (GPI) (Phosphoglucose isomerase) (PGI) (Phosphohexose
 isomerase) (PHI) (Fragment).
 DE Bos taurus (Bovine).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Savadye D.T.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE = D-FRUCTOSE 6-
 PHOSPHATE.
 CC -1- PATHWAY: INVOLVED IN GLYCOLYSIS AND IN GLUCONEOGENESIS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GPI FAMILY.
 DR EMBL; AF043228; AA97860.1; -.
 DR HSP; Q9NIE2; IHOX.
 DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
 DR GO; GO:0016853; F:isomerase activity; IEA.
 DR GO; GO:0006094; P:gluconeogenesis; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR001672; G6P_Isomerase.
 DR Pfam; PF00342; FGI; 1.
 DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
 KW Gluconeogenesis; Glycolysis; Isomerase.
 FT NON_TER 1 1
 FT NON_TER 85 85
 SQ SEQUENCE 85 AA; 9709 MW; 6E3ED81B3E0643B CRC64;

Query Match 92.4%; Score 85; DB 6; Length 85;
 Best Local Similarity 94.4%; Pred. No. 2.1e-06;
 Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHFEQLLSG 18
 |||||
 Db 51 LSIALHVGDFHFEQLLSG 68

RESULT 4
 ID Q95M65 PRELIMINARY; PRT; 186 AA.
 AC Q95M65;
 DT 01-DEC-2001 (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Autocrine motility factor (EC 5.3.1.9) (Glucose-6-phosphate isomerase)
 DE (GPI) (Phosphoglucose isomerase) (PGI) (Phosphohexose isomerase) (PHI)
 DE (Fragment).
 OS Mustela vison (American mink).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Mustelidae; Mustelinae;
 OC Mustela.
 OX NCBI_TaxID=9667;
 RN [1]

SEQUENCE FROM N.A.
 RA Bennett R.D., Murphy B.D.;
 RT "Autocrine motility factor and its receptor expression during
 implantation in the American mink.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE = D-FRUCTOSE 6-
 PHOSPHATE.
 CC -1- PATHWAY: INVOLVED IN GLYCOLYSIS AND IN GLUCONEOGENESIS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GPI FAMILY.
 DR EMBL; AY034381; AAK59814.1; -.
 DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
 DR GO; GO:0016853; F:isomerase activity; IEA.
 DR GO; GO:0006094; P:gluconeogenesis; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR001672; G6P_Isomerase.
 DR Pfam; PF00342; FGI; 1.
 DR PRINTS; PR00662; G6PISOMERASE.
 DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
 KW Gluconeogenesis; Glycolysis; Isomerase.
 FT NON_TER 1 1
 FT NON_TER 186 186
 SQ SEQUENCE 186 AA; 20606 MW; 68A950AESA9DD88 CRC64;

Query Match 92.4%; Score 85; DB 6; Length 186;
 Best Local Similarity 94.4%; Pred. No. 4.7e-06;

```

Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSIALHVGDFHFEQLLSG 18
Db 154 LSIALHVGDFNFEQLLSG 171
RESULT 5
QBHY2 PRELIMINARY; PRT; 281 AA.
AC Q8HY2;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Glucose-6-phosphate isomerase (Fragment).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Mustelidae; Mustelinae;
OC Mustela.
OX NCBI_TaxID=9668;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Corpus luteum;
RA Schulz L.C., Bahr J.M.;
RT "The Role of Glucose-6-Phosphate Isomerase in Embryo Implantation in
RL the Ferret (Mustela putorius).";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF154742; AA046115.1; -.
DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006094; P:gluconeogenesis; IEA.
DR GO; GO:0006096; P:glycolysis; IEA.
DR InterPro; IPR001672; G6P_Isomerase.
DR Pfam; PF00342; PGI; 1.
DR PRINTS; PR00662; G6PISOMERASE.
DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
KW Isomerase.
FT NON_TER 1
FT NON_TER 281
SQ SEQUENCE 281 AA; 31781 MW; E0FAFA2237607CB4 CRC64;

Query Match 92.4%; Score 85; DB 6; Length 281;
Best Local Similarity 94.4%; Pred. No. 7.2e-06;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSIALHVGDFHFEQLLSG 18
Db 51 LSIALHVGDFNFEQLLSG 68
RESULT 6
QBIRD3 PRELIMINARY; PRT; 300 AA.
AC Q9BRD3;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Body;
RA Strausberg R.;

```

```

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006342; AA006342.3; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 300 AA; 34160 MW; D1796B22A7B0BE54 CRC64;

Query Match 92.4%; Score 85; DB 4; Length 300;
Best Local Similarity 94.4%; Pred. No. 7.7e-06;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LSIALHVGDFHFEQLLSG 18
Db 24 LSIALHVGDFNFEQLLSG 41
RESULT 7
QBIRD3 PRELIMINARY; PRT; 553 AA.
AC Q7ZU30;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Glucose phosphate isomerase a.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Body;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Body;
RA Strausberg R.;

```

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC044450; AAH44450.1; -.
 DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
 DR GO; GO:0016853; F:isomerase activity; IEA.
 DR GO; GO:0006094; P:gluconeogenesis; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR001672; G6P_Isomerase.
 DR Pfam; PF00342; PGI; 1.
 DR PRINTS; PR00662; G6PISOMERASE.
 DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
 DR PROSITE; PS00174; P_GLUCOSE_ISOMERASE_2; 1.
 KW isomerase.
 SQ SEQUENCE 553 AA; 62072 MW; 6B18E750857A91D3 CRC64;

Query Match 91.3%; Score 84; DB 13; Length 553;
 Best Local Similarity 88.9%; Pred. No. 2.1e-05;
 Matches 16; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSTALHVGDFHFEQLLSG 18
 |||||:|||||
 Db 281 LSTALHVGDFHFEQLLSG 298
 |||||:|||||
 RESULT 8
 Q8QFT2 PRELIMINARY; PRT; 553 AA.
 AC Q8QFT2
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Phosphoglucose isomerase (EC 5.3.1.9) (Glucose-6-phosphate isomerase)
 DE (PGI) (Phosphohexose isomerase) (PHI).
 GN PGI-1.
 OS Mugil cephalus (Flathead mullet) (Mugil japonicus).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Mugilomorpha; Mugilidae;
 OC Mugil.
 OX NCBI_TaxID=48193;
 RN [1]
 RP MEDLINE=21918415; PubMed=11919278;
 RA Kao H.W., Lee S.C.;
 RT "Phosphoglucose isomerases of hagfish, zebrafish, grey mullet, toad,
 RT and snake, with reference to the evolution of the genes in
 RT vertebrates.";
 RL Mol. Biol. Evol. 19:367-374(2002).
 CC -|- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE = D-FRUCTOSE 6-PHOSPHATE.
 CC -|- PATHWAY: INVOLVED IN GLYCOLYSIS AND IN GLUCONEOGENESIS.
 CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -|- SIMILARITY: BELONGS TO THE GPI FAMILY.
 DR EMBL; AJ306392; CAC83778.1; -.
 DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
 DR GO; GO:0016853; F:isomerase activity; IEA.
 DR GO; GO:0006094; P:gluconeogenesis; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR001672; G6P_Isomerase.
 DR Pfam; PF00342; PGI; 1.
 DR PRINTS; PR00662; G6PISOMERASE.
 DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
 DR PROSITE; PS00174; P_GLUCOSE_ISOMERASE_2; 1.
 KW Gluconeogenesis; Glycolysis; Isomerase.
 SQ SEQUENCE 553 AA; 61963 MW; 09CGF6D8D07FF3EA CRC64;

Query Match 87.0%; Score 80; DB 13; Length 553;
 Best Local Similarity 83.3%; Pred. No. 9.8e-05;
 Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSTALHVGDFHFEQLLSG 18
 |||||:|||||
 Db 281 LSTALHVGDFHFEQLLSG 298
 |||||:|||||

RESULT 9
 Q8QFU2 PRELIMINARY; PRT; 553 AA.
 AC Q8QFU2
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Phosphoglucose isomerase-1 (EC 5.3.1.9) (Glucose-6-phosphate isomerase) (PGI) (Phosphohexose isomerase) (PHI).
 GN PGI-1.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP MEDLINE=21918415; PubMed=11919278;
 RA Kao H.W., Lee S.C.;
 RT "Phosphoglucose isomerases of hagfish, zebrafish, grey mullet, toad,
 RT and snake, with reference to the evolution of the genes in
 RT vertebrates.";
 RL Mol. Biol. Evol. 19:367-374(2002).
 CC -|- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE = D-FRUCTOSE 6-PHOSPHATE.
 CC -|- PATHWAY: INVOLVED IN GLYCOLYSIS AND IN GLUCONEOGENESIS.
 CC -|- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -|- SIMILARITY: BELONGS TO THE GPI FAMILY.
 DR EMBL; AJ306395; CAC83781.1; -.
 DR GO; GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
 DR GO; GO:0016853; F:isomerase activity; IEA.
 DR GO; GO:0006094; P:gluconeogenesis; IEA.
 DR GO; GO:0006096; P:glycolysis; IEA.
 DR InterPro; IPR001672; G6P_Isomerase.
 DR Pfam; PF00342; PGI; 1.
 DR PRINTS; PR00662; G6PISOMERASE.
 DR PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
 DR PROSITE; PS00174; P_GLUCOSE_ISOMERASE_2; 1.
 KW Gluconeogenesis; Glycolysis; Isomerase.
 SQ SEQUENCE 553 AA; 62084 MW; A1BE77B7D754C706 CRC64;

Query Match 87.0%; Score 80; DB 13; Length 553;
 Best Local Similarity 83.3%; Pred. No. 9.8e-05;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSTALHVGDFHFEQLLSG 18
 |||||:|||||
 Db 281 LSTALHVGDFHFEQLLSG 298
 |||||:|||||

RESULT 10
 Q8QFU6 PRELIMINARY; PRT; 553 AA.
 AC Q8QFU6
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Phosphoglucose isomerase (EC 5.3.1.9) (Glucose-6-phosphate isomerase) (PGI) (Phosphohexose isomerase) (PHI).
 GN Bufo melanostictus.
 OS Bufo melanostictus.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Bufo.
 OX NCBI_TaxID=30335;
 RN [1]
 RP MEDLINE=21918415; PubMed=11919278;
 RA Kao H.W., Lee S.C.;
 RT "Phosphoglucose isomerases of hagfish, zebrafish, grey mullet, toad,
 RT and snake, with reference to the evolution of the genes in
 RT vertebrates.";
 RL Mol. Biol. Evol. 19:367-374(2002).

Query Match 87.0%; Score 80; DB 13; Length 553;
 Best Local Similarity 83.3%; Pred. No. 9.8e-05;
 Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSTALHVGDFHFEQLLSG 18
 |||||:|||||
 Db 281 LSTALHVGDFHFEQLLSG 298
 |||||:|||||

CC -!- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE = D-FRUCTOSE 6-PHOSPHATE.
 CC -!- PATHWAY: INVOLVED IN GLYCOLYSIS AND IN GLUCONEOGENESIS.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE GPI FAMILY.
 DR EMBL: AJ308397; CAC83783.1; -.
 DR GO: GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
 DR GO: GO:0016853; F:isomerase activity; IEA.
 DR GO: GO:0006094; P:gluconeogenesis; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR001672; G6P_Isomerase.
 DR Pfam: PF00342; FGI; 1.
 DR PRINTS: PR00662; G6PISOMERASE.
 DR PROSITE: PS00765; P_GLUCOSE_ISOMERASE_1; 1.
 DR PROSITE: PS00174; P_GLUCOSE_ISOMERASE_2; 1.
 KW Gluconeogenesis; Glycolysis; Isomerase.
 SQ SEQUENCE 553 AA; 61737 MW; B0BBA201E781A16A CRC64;

Query Match 84.8%; Score 78; DB 13; Length 553;
 Best Local Similarity 83.3%; Pred. No. 0.00021;
 Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHFEQLLSG 18

Db 281 LSIALHVGDFHFEQLLAG 298

RESULT 11
 ID Q8QFU7 PRELIMINARY; PRT; 553 AA.
 AC Q8QFU7 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 24, Last annotation update)
 DE Phosphoglucose isomerase (EC 5.3.1.9) (Glucose-6-phosphate isomerase)
 DE (GPI) (PGI) (Phosphohexose isomerase) (PHI).
 GN Boiga krapelini.
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidodactylia; Squamata; Scleroglossa; Serpentes; Colubroidae;
 OC Colubridae; Colubrinae; Boiga.
 OX NCBI_TaxID=152266;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21918415; PubMed=11919278;
 RA Kao H.W., Lee S.C.;
 RT "Phosphoglucose isomerases of hagfish, zebrafish, grey mullet, toad, and snake, with reference to the evolution of the genes in vertebrates.";
 RL Mol. Biol. Evol. 19:367-374 (2002).
 CC -!- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE = D-FRUCTOSE 6-PHOSPHATE.
 CC -!- PATHWAY: INVOLVED IN GLYCOLYSIS AND IN GLUCONEOGENESIS.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE GPI FAMILY.
 DR EMBL: AJ308394; CAC83780.1; -.
 DR GO: GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
 DR GO: GO:0016853; F:isomerase activity; IEA.
 DR GO: GO:0006094; P:gluconeogenesis; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR001672; G6P_Isomerase.
 DR Pfam: PF00342; FGI; 1.
 DR PRINTS: PR00662; G6PISOMERASE.
 DR PROSITE: PS00765; P_GLUCOSE_ISOMERASE_1; 1.
 DR PROSITE: PS00174; P_GLUCOSE_ISOMERASE_2; 1.
 KW Gluconeogenesis; Glycolysis; Isomerase.
 SQ SEQUENCE 553 AA; 61978 MW; D4C81149E2A7E36 CRC64;

Query Match 83.7%; Score 77; DB 13; Length 553;
 Best Local Similarity 77.8%; Pred. No. 0.00031;
 Matches 14; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHFEQLLSG 18

Db 281 LSIALHVGDFHFEQLLAG 298

RESULT 12

ID Q8QFT1 PRELIMINARY; PRT; 553 AA.
 AC Q8QFT1 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 24, Last annotation update)
 DE Phosphoglucose isomerase-2 (EC 5.3.1.9) (Glucose-6-phosphate isomerase) (GPI) (PGI) (Phosphohexose isomerase) (PHI).
 DE (GPI) (PGI) (Phosphohexose isomerase) (PHI).
 GN Mugil cephalus (Flathead mullet) (Mugil japonicus).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Mugilomorpha; Mugilidae;
 OC Mugil.
 OX NCBI_TaxID=48193;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21918415; PubMed=11919278;
 RA Kao H.W., Lee S.C.;
 RT "Phosphoglucose isomerases of hagfish, zebrafish, grey mullet, toad, and snake, with reference to the evolution of the genes in vertebrates.";
 RL Mol. Biol. Evol. 19:367-374 (2002).
 CC -!- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE = D-FRUCTOSE 6-PHOSPHATE.
 CC -!- PATHWAY: INVOLVED IN GLYCOLYSIS AND IN GLUCONEOGENESIS.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE GPI FAMILY.
 DR EMBL: AJ306393; CAC83779.1; -.
 DR GO: GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
 DR GO: GO:0016853; F:isomerase activity; IEA.
 DR GO: GO:0006094; P:gluconeogenesis; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR001672; G6P_Isomerase.
 DR Pfam: PF00342; FGI; 1.
 DR PRINTS: PR00662; G6PISOMERASE.
 DR PROSITE: PS00765; P_GLUCOSE_ISOMERASE_1; 1.
 DR PROSITE: PS00174; P_GLUCOSE_ISOMERASE_2; 1.
 KW Gluconeogenesis; Glycolysis; Isomerase.
 SQ SEQUENCE 553 AA; 62147 MW; E5F0F8B5F8C6D9A CRC64;

Query Match 81.5%; Score 75; DB 13; Length 553;
 Best Local Similarity 72.2%; Pred. No. 0.00066;
 Matches 13; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHFEQLLSG 18

Db 281 MAIALHVGDFHFEQLSG 298

RESULT 13
 ID Q8T313 PRELIMINARY; PRT; 554 AA.
 AC Q8T313 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 24, Last annotation update)
 DE Phosphoglucose isomerase (EC 5.3.1.9) (Glucose-6-phosphate isomerase)
 DE (GPI) (PGI) (Phosphohexose isomerase) (PHI).
 GN Paramyxine yangi.
 OS Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
 OC Myxiniidae; Eptatretinae; Paramyxine.
 OX NCBI_TaxID=94968;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21918415; PubMed=11919278;
 RA Kao H.W., Lee S.C.;

Query Match 81.5%; Score 75; DB 13; Length 553;
 Best Local Similarity 72.2%; Pred. No. 0.00066;
 Matches 13; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHFEQLLSG 18

Db 281 MAIALHVGDFHFEQLSG 298

RESULT 13
 ID Q8T313 PRELIMINARY; PRT; 554 AA.
 AC Q8T313 (TRENBLrel. 21, Created)
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 24, Last annotation update)
 DE Phosphoglucose isomerase (EC 5.3.1.9) (Glucose-6-phosphate isomerase)
 DE (GPI) (PGI) (Phosphohexose isomerase) (PHI).
 GN Paramyxine yangi.
 OS Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
 OC Myxiniidae; Eptatretinae; Paramyxine.
 OX NCBI_TaxID=94968;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21918415; PubMed=11919278;
 RA Kao H.W., Lee S.C.;

RT "Phosphoglucose isomerases of hagfish, zebrafish, grey mullet, toad,
 RT and snake, with reference to the evolution of the genes in
 RT vertebrates";
 RL Mol. Biol. Evol. 19:367-374(2002).
 CC -1- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE = D-FRUCTOSE 6-
 CC PHOSPHATE.
 CC -1- PATHWAY: INVOLVED IN GLYCOLYSIS AND IN GLUCONEOGENESIS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GPI FAMILY.
 DR EMBL: AJ306391; CAC83777.1; -
 DR GO: GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
 DR GO: GO:0016853; F:isomerase activity; IEA.
 DR GO: GO:0006094; P:gluconeogenesis; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR001672; G6P_Isomerase.
 DR Pfam: PF00342; PGI.1.
 DR PRINTS: PR00662; G6PISOMERASE.
 DR PROSITE: PS00765; P_GLUCOSE_ISOMERASE_1; 1.
 DR PROSITE: PS00174; P_GLUCOSE_ISOMERASE_2; 1.
 KW Gluconeogenesis; Glycolysis; Isomerase.
 SQ SEQUENCE 554 AA; 62031 MW; 6BD9A08351A30683 CRC64;

Query Match 77.2%; Score 71; DB 5; Length 554;
 Best Local Similarity 72.2%; Pred. No. 0.0031;
 Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSIALHVGFDHFQQLSG 18
 |||||:|:|:|:|:|
 Db 282 LSIALHGFENFELQLG 299

RESULT 14

Q9UIQ2 ID Q9UIQ2 PRELIMINARY; PRT; 551 AA.
 AC Q9UIQ2
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Y87G2A.8 protein (EC 5.3.1.9) (Glucose-6-phosphate isomerase) (GPI)
 DE (Phosphoglucose isomerase) (PGI) (Phosphohexose isomerase)
 DE (PHI).
 GN Y87G2A.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA White S.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology";
 RL Science 282:2012-2018(1998).
 CC -1- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE = D-FRUCTOSE 6-
 CC PHOSPHATE.
 CC -1- PATHWAY: INVOLVED IN GLYCOLYSIS AND IN GLUCONEOGENESIS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GPI FAMILY.
 DR EMBL: AL110500; CAB60430.1; -
 DR HSSP: Q9N1E2; IHOX.
 DR WormPep: Y87G2A.8; CE24687.
 DR GO: GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
 DR GO: GO:0016853; F:isomerase activity; IEA.
 DR GO: GO:0006094; P:gluconeogenesis; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR001672; G6P_Isomerase.
 DR Pfam: PF00342; PGI.1.
 DR PRINTS: PR00662; G6PISOMERASE.
 DR PROSITE: PS00765; P_GLUCOSE_ISOMERASE_1; 1.

DR PROSITE: PS00174; P_GLUCOSE_ISOMERASE_2; 1.
 KW Gluconeogenesis; Glycolysis; Isomerase.
 SQ SEQUENCE 551 AA; 61122 MW; A2BF1E1E13F1916C CRC64;

Query Match 76.1%; Score 70; DB 5; Length 551;
 Best Local Similarity 66.7%; Pred. No. 0.0045;
 Matches 12; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSIALHVGFDHFQQLSG 18
 |||||:|:|:|:|:|
 Db 280 LSIAVHIGFDNYEKLDG 297

RESULT 15

Q8QFUL ID Q8QFUL PRELIMINARY; PRT; 553 AA.
 AC Q8QFUL
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Phosphoglucose isomerase-2 (EC 5.3.1.9) (Glucose-6-phosphate
 DE isomerase) (GPI) (PGI) (Phosphohexose isomerase) (PHI).
 GN PGI-2.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=21918415; PubMed=11919278;
 RA Kao H.W., Lee S.C.;
 RT "Phosphoglucose isomerases of hagfish, zebrafish, grey mullet, toad,
 RT and snake, with reference to the evolution of the genes in
 RT vertebrates";
 RL Mol. Biol. Evol. 19:367-374(2002).
 CC -1- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE = D-FRUCTOSE 6-
 CC PHOSPHATE.
 CC -1- PATHWAY: INVOLVED IN GLYCOLYSIS AND IN GLUCONEOGENESIS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GPI FAMILY.
 DR EMBL: AJ306396; CAC83782.1; -
 DR GO: GO:0004347; F:glucose-6-phosphate isomerase activity; IEA.
 DR GO: GO:0016853; F:isomerase activity; IEA.
 DR GO: GO:0006094; P:gluconeogenesis; IEA.
 DR GO: GO:0006096; P:glycolysis; IEA.
 DR InterPro: IPR001672; G6P_Isomerase.
 DR Pfam: PF00342; PGI.1.
 DR PRINTS: PR00662; G6PISOMERASE.
 DR PROSITE: PS00765; P_GLUCOSE_ISOMERASE_1; 1.
 DR PROSITE: PS00174; P_GLUCOSE_ISOMERASE_2; 1.
 KW Gluconeogenesis; Glycolysis; Isomerase.
 SQ SEQUENCE 553 AA; 62096 MW; EB7EC7A244442D47 CRC64;

Query Match 76.1%; Score 70; DB 13; Length 553;
 Best Local Similarity 66.7%; Pred. No. 0.0045;
 Matches 12; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSIALHVGFDHFQQLSG 18
 |||||:|:|:|:|:|
 Db 281 LSIALHIGYENFEKLLAG 298

Search completed: May 19, 2004, 15:46:26
 Job time : 38.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 15:27:51 ; Search time 40.5 Seconds
(without alignments)
123.672 Million cell updates/sec

Title: US-10-005-684-7

Perfect score: 92
Sequence: 1 LSTALHVGDFHFEQLLSG 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1145568 seqs, 278261457 residues

Total number of hits satisfying chosen parameters: 1145568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCF_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCFUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 92 | 100.0 | 18 | 14 | US-10-005-684-7 |
| 2 | 85 | 92.4 | 558 | 12 | US-10-170-385-367 |
| 3 | 85 | 92.4 | 558 | 14 | US-10-091-333-8 |
| 4 | 85 | 92.4 | 558 | 14 | US-10-325-878-8 |
| 5 | 85 | 92.4 | 558 | 14 | US-10-177-293-188 |
| 6 | 79 | 85.9 | 548 | 12 | US-10-282-122A-78480 |
| 7 | 70 | 76.1 | 551 | 14 | US-10-179-766-5 |
| 8 | 69 | 75.0 | 549 | 12 | US-10-282-122A-66993 |
| 9 | 68 | 73.9 | 548 | 12 | US-10-282-122A-68616 |
| 10 | 68 | 73.9 | 548 | 15 | US-10-369-493-21211 |
| 11 | 65 | 70.7 | 548 | 15 | US-10-369-493-504 |
| 12 | 65 | 70.7 | 553 | 14 | US-10-128-714-3098 |
| 13 | 65 | 70.7 | 553 | 14 | US-10-128-714-8098 |
| 14 | 62 | 67.4 | 509 | 15 | US-10-369-493-12559 |
| 15 | 62 | 67.4 | 563 | 12 | US-10-282-122A-58546 |

| | | | | | | |
|----|----|------|------|----|----------------------|-------------------|
| 16 | 62 | 67.4 | 592 | 9 | US-09-934-901-14 | Sequence 14, Appl |
| 17 | 62 | 67.4 | 592 | 9 | US-09-934-868-4 | Sequence 4, Appl |
| 18 | 62 | 67.4 | 592 | 14 | US-10-320-924-14 | Sequence 14, Appl |
| 19 | 62 | 67.4 | 592 | 14 | US-10-320-874-14 | Sequence 14, Appl |
| 20 | 62 | 67.4 | 658 | 15 | US-10-320-797-3267 | Sequence 3267, Ap |
| 21 | 61 | 66.3 | 533 | 15 | US-10-369-493-8472 | Sequence 8472, Ap |
| 22 | 61 | 66.3 | 542 | 12 | US-10-282-122A-56173 | Sequence 56173, A |
| 23 | 61 | 66.3 | 542 | 15 | US-10-369-493-10367 | Sequence 10367, A |
| 24 | 61 | 66.3 | 546 | 12 | US-10-282-122A-73359 | Sequence 73359, A |
| 25 | 61 | 66.3 | 549 | 12 | US-10-282-122A-43263 | Sequence 43263, A |
| 26 | 61 | 66.3 | 549 | 12 | US-10-282-122A-59757 | Sequence 59757, A |
| 27 | 61 | 66.3 | 549 | 12 | US-10-282-122A-75162 | Sequence 75162, A |
| 28 | 61 | 66.3 | 549 | 12 | US-10-282-122A-75758 | Sequence 75758, A |
| 29 | 61 | 66.3 | 549 | 15 | US-10-369-493-23655 | Sequence 23655, A |
| 30 | 60 | 65.2 | 550 | 15 | US-10-369-493-22572 | Sequence 22572, A |
| 31 | 60 | 65.2 | 554 | 15 | US-10-369-493-1484 | Sequence 1484, Ap |
| 32 | 57 | 62.0 | 500 | 15 | US-10-369-493-15351 | Sequence 15351, A |
| 33 | 57 | 62.0 | 551 | 15 | US-10-369-493-16105 | Sequence 16105, A |
| 34 | 57 | 62.0 | 562 | 15 | US-10-369-493-15718 | Sequence 15718, A |
| 35 | 55 | 59.8 | 580 | 15 | US-10-369-493-4061 | Sequence 4061, Ap |
| 36 | 54 | 58.7 | 538 | 15 | US-10-369-493-19717 | Sequence 19717, A |
| 37 | 54 | 58.7 | 550 | 12 | US-10-282-122A-77028 | Sequence 77028, A |
| 38 | 53 | 57.6 | 550 | 14 | US-10-032-585-7207 | Sequence 7207, Ap |
| 39 | 51 | 55.4 | 556 | 12 | US-10-282-122A-45095 | Sequence 45095, A |
| 40 | 50 | 54.3 | 539 | 15 | US-10-369-493-13725 | Sequence 13725, A |
| 41 | 50 | 54.3 | 554 | 12 | US-10-282-122A-69270 | Sequence 69270, A |
| 42 | 49 | 53.3 | 554 | 12 | US-10-282-122A-66717 | Sequence 66717, A |
| 43 | 48 | 52.2 | 1230 | 14 | US-10-156-761-14097 | Sequence 14097, A |
| 44 | 46 | 50.0 | 554 | 12 | US-10-282-122A-68068 | Sequence 68068, A |
| 45 | 46 | 50.0 | 554 | 12 | US-10-282-122A-68085 | Sequence 68085, A |

ALIGNMENTS

RESULT 1
US-10-005-684-7
; Sequence 7, Application US/10005684
; Publication No. US20030100035A1
; GENERAL INFORMATION:
; APPLICANT: Vojdani, Aristo
; TITLE OF INVENTION: SALIVA IMMUNOASSAY FOR DETECTION OF
; FILE REFERENCE: IMSC12.005A
; CURRENT APPLICATION NUMBER: US/10/005,684
; CURRENT FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Arthritis peptide
US-10-005-684-7

Query Match 100.0%; Score 92; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSTALHVGDFHFEQLLSG 18
Db 1 LSTALHVGDFHFEQLLSG 18

RESULT 2
US-10-170-385-367
; Sequence 367, Application US/10170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan

APPLICANT: Binley, Katie Mary
APPLICANT: Rayner, William Nigel
APPLICANT: Naylor, Stuart
APPLICANT: Kingman, Susan Mary
APPLICANT: Krige, David
TITLE OF INVENTION: ANALYSIS METHOD
FILE REFERENCE: 53268200100
CURRENT APPLICATION NUMBER: US/10/170,385
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: PCT/GB02/01662
PRIOR FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: PCT/GB01/05458
PRIOR FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 549
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 367
LENGTH: 558
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-170-385-367

Query Match 92.4%; Score 85; DB 12; Length 558;
Best Local Similarity 94.4%; Pred. No. 7.9e-06;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSIALHVGDFHFEQLLSG 18
|||||:|||||
Db 282 LSIALHVGDFNFEQLLSG 299

RESULT 3
US-10-091-333-8
Sequence 8, Application US/10091333
Publication No. US20030104973A1
GENERAL INFORMATION:
APPLICANT: EINAT, Paz
APPLICANT: SKALITER, Rami
TITLE OF INVENTION: HYPOXIA-REGULATED GENES
FILE REFERENCE: EINAT-1.1D
CURRENT APPLICATION NUMBER: US/10/091,333
CURRENT FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: US 09/604,978
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: US 09/138,112
PRIOR FILING DATE: 1999-08-21
PRIOR APPLICATION NUMBER: US 60/056,453
PRIOR FILING DATE: 1997-08-21
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8
LENGTH: 558
TYPE: PRT
ORGANISM: Rattus norvegicus
US-10-091-333-8

Query Match 92.4%; Score 85; DB 14; Length 558;
Best Local Similarity 94.4%; Pred. No. 7.9e-06;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSIALHVGDFHFEQLLSG 18
|||||:|||||
Db 282 LSIALHVGDFNFEQLLSG 299

RESULT 4
US-10-325-878-8
Sequence 8, Application US/10325878
Publication No. US20030124118A1
GENERAL INFORMATION:
APPLICANT: EINAT, Paz
APPLICANT: SKALITER, Rami
TITLE OF INVENTION: HYPOXIA-REGULATED GENES
NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:
ADDRESSEE: KOHN & ASSOCIATES
STREET: 30500 NO. US20030124116Althwestern Hwy., Suite 401
CITY: Farmington Hills
STATE: Michigan
COUNTRY: U.S.
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/325,878
FILING DATE: 23-Dec-2002
CLASSIFICATION: <UNKNOWN>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/138,112
FILING DATE: <UNKNOWN>
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 0168.00034
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 5395055
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 558 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-325-878-8

Query Match 92.4%; Score 85; DB 14; Length 558;
Best Local Similarity 94.4%; Pred. No. 7.9e-06;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSIALHVGDFHFEQLLSG 18
|||||:|||||
Db 282 LSIALHVGDFNFEQLLSG 299

RESULT 5
US-10-177-293-188
Sequence 188, Application US/10177293
Publication No. US20030124128A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Glatt, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Gannavarpu, Manjula
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Mertens, Maureen
APPLICANT: Myer, Vic
APPLICANT: Wang, Youzhen
APPLICANT: Xu, Yongyao
APPLICANT: Hoersch, Sebastian
APPLICANT: Monahan, John
APPLICANT: Meyers, Rachel E.
APPLICANT: Bast Jr., Robert C.
APPLICANT: Hortobagyi, Gabriel N.
APPLICANT: Pusztai, Lajos
APPLICANT: Meric, Funda
APPLICANT: Sahin, Aysegul
APPLICANT: Mills, Gordon B.
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
PREVENTION, AND THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-038
CURRENT APPLICATION NUMBER: US/10/177,293

```
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 188
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-188

Query Match          92.4%; Score 85; DB 14; Length 558;
Best Local Similarity 94.4%; Pred. No. 7.9e-06;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHPEQLLSG 18
      |||||:|||||:|||||
Db 282 LSIALHVGDFHPEQLLSG 299

RESULT 6
US-10-282-122A-78480
; Sequence 78480, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78480
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Yersinia pestis
US-10-282-122A-78480

Query Match          85.9%; Score 79; DB 12; Length 548;
Best Local Similarity 88.9%; Pred. No. 8.1e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHPEQLLSG 18
      |||||:|||||:|||||
Db 279 LSIALSVGFHPEQLLSG 296

RESULT 7
US-10-179-766-5
; Sequence 5, Application US/10179766
; Publication No. US20030190312A1
; GENERAL INFORMATION:
; APPLICANT: Kenyon, Cynthia
; APPLICANT: Apfeld, Javier
; APPLICANT: Dillin, Andrew
; APPLICANT: Garigan, Delia
; APPLICANT: Hsu, Ao-Lin A.
; APPLICANT: Lehrer-Graiwer, Josh
; APPLICANT: Murphy, Coleen
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Eukaryotic Genes Involved in Adult Lifespan Regulation
; FILE REFERENCE: 023070-119930US
; CURRENT APPLICATION NUMBER: US/10/179,766
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: US 60/300,577
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/301,052
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/373,975
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; OTHER INFORMATION: glucose phosphate isomerase (phosphoglucose
; isomerase), gene Y87G2A.8 (GPI-1)
US-10-179-766-5

Query Match          76.1%; Score 70; DB 14; Length 551;
Best Local Similarity 66.7%; Pred. No. 0.0027;
Matches 12; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHPEQLLSG 18
      |||||:|||||:|||||
Db 280 LSIATVHGFNDYKLLDG 297

RESULT 8
US-10-282-122A-66993
; Sequence 66993, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
```

```
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 66993
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-282-122A-66993

Query Match 75.0%; Score 69; DB 12; Length 549;
Best Local Similarity 72.2%; Pred. No. 0.004;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LSIALHVGFDHFEQLLSG 18
Db 279 LSIALSIGFHFALLAG 296

RESULT 9
US-10-282-122A-68616
; Sequence 68616, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
```

```
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 68616
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-10-282-122A-68616

Query Match 73.9%; Score 68; DB 12; Length 548;
Best Local Similarity 72.2%; Pred. No. 0.0059;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LSIALHVGFDHFEQLLSG 18
Db 279 LSIVLSIGYDNFEQLLSG 296

RESULT 10
US-10-369-493-21211
; Sequence 21211, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21211
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-21211

Query Match 73.9%; Score 68; DB 15; Length 548;
Best Local Similarity 72.2%; Pred. No. 0.0059;
Matches 13; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LSIALHVGFDHFEQLLSG 18
Db 279 LSIALSIGYDNFEQLLSG 296

RESULT 11
US-10-369-493-504
; Sequence 504, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
```

```

? LOCATION: (11)::(1309)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-12559

Query Match          Score 62;   DB 15;   Length 509;
Best Local Similarity 55.6%;   Pred. No. 0.056;
Matches 10;   Conservative 6;   Mismatches 2;   Indels 0;   Gaps 0;

Qy      1  LSIALHVGFDHFFQLLSG 18
        ||::||::||::||::|
Db      243  LSVALYIGYDNFHQFLAG 260

```

```

RESULT 15
US-10-282-122A-58546
; Sequence 58546, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58546
; LENGTH: 563
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-282-122A-58546

Query Match      67.4%; Score 62; DB 12; Length 563;
Best Local Similarity 66.7%; Pred. No. 0.063;
Matches 12; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHFEQLLSG 18
    ||||| :||:| |
Db 293 LSIALSIGFENFEALING 310

Search completed: May 19, 2004, 15:41:17
Job time : 40.5 secs

```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 15:23:26 ; Search time 15.75 Seconds
(without alignments)
59.001 Million cell updates/sec

Title: US-10-005-684-7
Perfect score: 92
Sequence: 1 LSIALHVGDFHFEQLISG 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pcp.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 85 | 92.4 | 558 | 4 | US-09-604-978-8 |
| 2 | 85 | 92.4 | 558 | 4 | US-09-604-728-8 |
| 3 | 88 | 73.9 | 548 | 4 | US-09-343-681A-7327 |
| 4 | 62 | 67.4 | 592 | 4 | US-09-934-901-14 |
| 5 | 61 | 66.3 | 593 | 4 | US-09-489-039A-8609 |
| 6 | 51 | 55.4 | 565 | 4 | US-09-328-352-4920 |
| 7 | 49 | 53.3 | 565 | 4 | US-09-328-352-2702 |
| 8 | 45 | 48.9 | 298 | 4 | US-09-328-352-6225 |
| 9 | 43.5 | 47.3 | 616 | 4 | US-09-489-039A-9231 |
| 10 | 43 | 46.7 | 103 | 4 | US-09-328-352-5690 |
| 11 | 42 | 45.7 | 538 | 4 | US-09-252-991A-17359 |
| 12 | 42 | 45.7 | 573 | 4 | US-09-540-236-2187 |
| 13 | 41 | 44.6 | 118 | 4 | US-09-765-815-8 |
| 14 | 41 | 44.6 | 580 | 1 | US-08-420-235B-15 |
| 15 | 41 | 44.6 | 580 | 3 | US-08-793-624-15 |
| 16 | 41 | 44.6 | 580 | 5 | PCT-US95-10194-15 |
| 17 | 41 | 44.6 | 591 | 3 | US-09-082-737-2 |
| 18 | 41 | 44.6 | 591 | 4 | US-09-688-188B-103 |
| 19 | 41 | 44.6 | 591 | 4 | US-09-718-032-2 |
| 20 | 41 | 44.6 | 591 | 4 | US-09-291-417D-103 |
| 21 | 40.5 | 44.0 | 321 | 2 | US-08-592-126-143 |
| 22 | 40.5 | 44.0 | 321 | 4 | US-09-168-595-143 |
| 23 | 40.5 | 44.0 | 420 | 2 | US-08-592-126-142 |
| 24 | 40.5 | 44.0 | 420 | 4 | US-09-168-595-142 |
| 25 | 40 | 43.5 | 95 | 4 | US-09-328-352-5647 |
| 26 | 40 | 43.5 | 1043 | 4 | US-09-252-991A-28885 |
| 27 | 39 | 42.4 | 169 | 4 | US-09-107-532A-7306 |

| | | | | | | |
|----|------|------|------|---|----------------------|--------------------|
| 28 | 39 | 42.4 | 204 | 4 | US-09-199-637A-225 | Sequence 225, App |
| 29 | 39 | 42.4 | 267 | 4 | US-09-489-039A-10578 | Sequence 10578, A |
| 30 | 39 | 42.4 | 291 | 4 | US-09-495-406-3 | Sequence 3, Appli |
| 31 | 39 | 42.4 | 291 | 4 | US-09-495-406-5 | Sequence 5, Appli |
| 32 | 39 | 42.4 | 291 | 4 | US-09-495-406-7 | Sequence 7, Appli |
| 33 | 39 | 42.4 | 291 | 4 | US-09-495-406-9 | Sequence 9, Appli |
| 34 | 39 | 42.4 | 548 | 6 | 5256558-5 | Patent No. 5256558 |
| 35 | 39 | 42.4 | 1178 | 1 | US-08-199-776-2 | Sequence 2, Appli |
| 36 | 39 | 42.4 | 1178 | 3 | US-08-663-731-2 | Sequence 2, Appli |
| 37 | 39 | 42.4 | 1178 | 3 | US-08-879-338-2 | Sequence 2, Appli |
| 38 | 39 | 42.4 | 1178 | 5 | PCT-US95-02044-2 | Sequence 2, Appli |
| 39 | 39 | 42.4 | 1179 | 4 | US-09-293-238B-2 | Sequence 2, Appli |
| 40 | 38.5 | 41.8 | 425 | 2 | US-08-951-148-3 | Sequence 3, Appli |
| 41 | 38.5 | 41.8 | 425 | 2 | US-09-165-234-3 | Sequence 3, Appli |
| 42 | 38.5 | 41.8 | 425 | 3 | US-09-274-570-3 | Sequence 3, Appli |
| 43 | 38 | 41.3 | 112 | 1 | US-08-446-919A-5 | Sequence 5, Appli |
| 44 | 38 | 41.3 | 205 | 1 | US-09-540-236-3658 | Sequence 3658, Ap |
| 45 | 38 | 41.3 | 286 | 4 | US-09-489-039A-9521 | Sequence 9521, Ap |

ALIGNMENTS

RESULT 1
US-09-604-978-8
; Sequence 8, Application US/09604978
; Patent No. 6455674
; GENERAL INFORMATION:
; APPLICANT: Einat, Paz
; TITLE OF INVENTION: HYPOXIA-REGULATED GENES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6455674thwestern Hwy., Suite 401
; City: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48134
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/604,978
; FILING DATE: 28-Jun-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/138,112
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 0168.00034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 5395055
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 558 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-604-978-8

Query Match 92.4%; Score 85; DB 4; Length 558;
Best Local Similarity 94.4%; Pred. No. 4e-07;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSIALHVGDFHFEQLLSG 18
| | | | | | | | | | | | | | | | | | | | | |
Db 282 LSIALHVGDFHFEQLLSG 299

RESULT 2
US-09-604-728-8
; Sequence 8, Application US/09604728
; Patent No. 6555667
; GENERAL INFORMATION:
; APPLICANT: Einat, Paz
; TITLE OF INVENTION: HYPOXIA-REGULATED GENES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6555667thwestern Hwy., Suite 401
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/604,728
; FILING DATE: 28-Jun-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/138,112
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 0168.00034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 558 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-09-604-728-8
Query Match 92.4%; Score 85; DB 4; Length 558;
Best Local Similarity 94.4%; Pred. No. 4e-07;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSIALHVGDFHFEQLLSG 18
| | | | | | | | | | | | | | | | | | | | | |
Db 282 LSIALHVGDFHFEQLLSG 299

RESULT 3
US-09-543-681A-7327
; Sequence 7327, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7327
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7327

Query Match 73.9%; Score 68; DB 4; Length 548;
Best Local Similarity 72.2%; Pred. No. 0.00045;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LSIALHVGDFHFEQLLSG 18
| | | | | | | | | | | | | | | | | | | | | |
Db 279 LSIVLSIGYDNFEQLLSG 296

RESULT 4
US-09-934-901-14
; Sequence 14, Application US/09934901
; Patent No. 6555353
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, J. Martin
; APPLICANT: No. 6555353ton, Kelley C.
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: DENTRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: C11619 US NA
; CURRENT APPLICATION NUMBER: US/09/934,901
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,906
; PRIOR FILING DATE: September 1, 2000
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 592
; TYPE: PRT
; ORGANISM: METHYLOMONAS SP.
US-09-934-901-14

Query Match 67.4%; Score 62; DB 4; Length 592;
Best Local Similarity 61.1%; Pred. No. 0.0059;
Matches 11; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LSIALHVGDFHFEQLLSG 18
| | | | | | | | | | | | | | | | | | | | | |
Db 324 MSIALYIGMDNFELLNG 341

RESULT 5
US-09-489-039A-8609
; Sequence 8609, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8609
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8609

Query Match 66.3%; Score 61; DB 4; Length 593;
Best Local Similarity 72.2%; Pred. No. 0.0089;
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LSIALHVGDFHFEQLLSG 18

Db 323 LSIILSVGDFNFVELLSG 340
||||| :|||

RESULT 6

US-09-328-352-4920
; Sequence 4920, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: GTC99-03PA
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4920
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4920

Query Match 55.4%; Score 51; DB 4; Length 565;
Best Local Similarity 62.5%; Pred. No. 0.53;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 IALHVGDFHFEQLLSG 18
||| :|||

Db 296 IALTIGDGFQQLLAG 311
||| :|||

RESULT 7

US-09-252-991A-27022
; Sequence 27022, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: 107196.136
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27022
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27022

Query Match 53.3%; Score 49; DB 4; Length 565;
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LSIALHVGDFHFEQLLSG 18
||| :|||

Db 294 LPIAMSGISNFKELLSG 311
||| :|||

RESULT 8

US-09-328-352-6225
; Sequence 6225, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: GTC99-03PA
; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6225
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6225

Query Match 48.9%; Score 45; DB 4; Length 298;
Best Local Similarity 50.0%; Pred. No. 3;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LSIALHVGDFHFEQLLSG 18
||| :|||

Db 134 LQALTCAPFHFTALLGG 151
||| :|||

RESULT 9

US-09-489-039A-9231
; Sequence 9231, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: 2709.2004001
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9231
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9231

Query Match 47.3%; Score 43.5; DB 4; Length 616;
Best Local Similarity 62.5%; Pred. No. 13;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 1 LSIALHVGDFHFEQLL 16
||| :|||

Db 282 LQAALHI-FFHFOQLL 296
||| :|||

RESULT 10

US-09-328-352-5690
; Sequence 5690, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: GTC99-03PA
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5690
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5690

Query Match 46.7%; Score 43; DB 4; Length 103;
Best Local Similarity 56.2%; Pred. No. 2;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 SIALHVGDFHFEQLLS 17
||| :|||

Db 23 SITIHPTDKHQQLAS 38
||| :|||

RESULT 11

```
US-09-252-991A-17359
; Sequence 17359, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17359
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17359
Query Match 45.7%; Score 42; DB 4; Length 538;
Best Local Similarity 44.4%; Pred. No. 21;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LSIALHVGDFHFEQLLSG 18
   : : : : :
Db 88 LQVGEHLILDHYAQLLVG 105

RESULT 12
US-09-540-236-2187
; Sequence 2187, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2187
; LENGTH: 573
; TYPE: PRT
; ORGANISM: M. catarrhalis
US-09-540-236-2187
Query Match 45.7%; Score 42; DB 4; Length 573;
Best Local Similarity 44.4%; Pred. No. 23;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LSIALHVGDFHFEQLLSG 18
   : : : : :
Db 301 LAIAVRIGMANFRALLAG 318

RESULT 13
US-09-765-815-8
; Sequence 8, Application US/09765815
; Patent No. 6673586
; GENERAL INFORMATION:
; APPLICANT: Balk, Steven
; TITLE OF INVENTION: No. 6673586el Steroid Hormone Receptor
; FILE REFERENCE: 01948/068002
; CURRENT APPLICATION NUMBER: US/09/765,815
; CURRENT FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/176,859
; PRIOR FILING DATE: 2000-01-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
```

```
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-765-815-8
Query Match 44.6%; Score 41; DB 4; Length 118;
Best Local Similarity 50.0%; Pred. No. 5.4;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 5 LHVGDFHFEQLLSG 18
   : : : : :
Db 21 VHTGFDQHEQKFTG 34

RESULT 14
US-08-420-235B-15
; Sequence 15, Application US/08420235B
; Patent No. 5801042
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/420,235B
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 45185-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 580 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-420-235B-15
Query Match 44.6%; Score 41; DB 1; Length 580;
Best Local Similarity 72.7%; Pred. No. 35;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 VGFDHFEQLLS 17
   : : : : :
Db 362 LSPDHPFFQLLS 392

RESULT 15
US-08-793-624-15
; Sequence 15, Application US/08793624C
; Patent No. 6150093
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Unique Associated Kaposi's Sarcoma Virus Sequences And
; TITLE OF INVENTION: Uses Thereof
```

```

; FILE REFERENCE: 45185-C-PCT-US/JPW
; CURRENT APPLICATION NUMBER: US/08/793,624C
; CURRENT FILING DATE: 1997-02-18
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-08-793-624-15

Query Match      44.6%; Score 41; DB 3; Length 580;
Best Local Similarity 72.7%; Pred. No. 35;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      7 VGFDFHFEQLLS 17
Db      : |||||
        382 LSFDFHFEQLLS 392

Search completed: May 19, 2004, 15:29:41
Job time : 16.75 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: May 19, 2004, 15:38:37 ; Search time 41.6667 Seconds
(without alignments)
151.449 Million cell updates/sec

Title: US-10-005-684-6
Perfect score: 108
Sequence: 1 KARIHPFHILIALETYKTH 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phase: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_rvirus: *
16: sp_bacteriap: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|---------------------|
| 1 | 108 | 100.0 | 518 | 4 Q86WL4 | Q86wl4 homo sapien |
| 2 | 108 | 100.0 | 534 | 4 Q86WL3 | Q86wl3 homo sapien |
| 3 | 108 | 100.0 | 538 | 4 Q9H1W6 | Q9hlw6 homo sapien |
| 4 | 99 | 91.7 | 538 | 11 Q8R562 | Q8r562 mus musculus |
| 5 | 99 | 91.7 | 538 | 11 Q8OW12 | Q8ow12 mus musculus |
| 6 | 99 | 91.7 | 547 | 11 Q7TT40 | Q7tt40 mus musculus |
| 7 | 63 | 58.3 | 552 | 9 Q852Y5 | Q852y5 mycobacteri |
| 8 | 48 | 44.4 | 385 | 16 Q81P65 | Q81p65 bacillus an |
| 9 | 48 | 44.4 | 997 | 5 Q01858 | Q01858 caenorhabdi |
| 10 | 47 | 43.5 | 188 | 10 Q9AW09 | Q9aw09 guillardi |
| 11 | 47 | 43.5 | 219 | 5 Q8ILV2 | Q8ilv2 plasmodium |
| 12 | 47 | 43.5 | 266 | 2 Q33728 | Q33728 streptococ |
| 13 | 46 | 42.6 | 173 | 10 Q40650 | Q40650 oryza sativ |
| 14 | 46 | 42.6 | 177 | 10 Q7X9L9 | Q7x9l9 triticum ae |
| 15 | 46 | 42.6 | 210 | 5 Q8MTT4 | Q8mtt4 entamoeba h |
| 16 | 46 | 42.6 | 224 | 10 Q9FUN3 | Q9fun3 elaeis guin |

| | | | | | |
|----|------|------|------|-----------|---------------------|
| 17 | 45 | 41.7 | 134 | 16 Q89XM3 | Q89xm3 bradyrhizob |
| 18 | 45 | 41.7 | 150 | 11 Q8CD31 | Q8cd31 mus musculus |
| 19 | 45 | 41.7 | 184 | 10 Q9C6H7 | Q9c6h7 arabidopsis |
| 20 | 45 | 41.7 | 220 | 10 Q93VT9 | Q93vt9 arabidopsis |
| 21 | 45 | 41.7 | 221 | 10 Q93W22 | Q93w22 arabidopsis |
| 22 | 45 | 41.7 | 222 | 5 Q7YWM1 | Q7ywm1 cryptospori |
| 23 | 45 | 41.7 | 338 | 3 Q74426 | Q74426 schizosacch |
| 24 | 45 | 41.7 | 475 | 13 Q98U17 | Q98u17 alcolapia g |
| 25 | 45 | 41.7 | 475 | 13 Q9PU05 | Q9puq5 opsanus bet |
| 26 | 45 | 41.7 | 486 | 13 Q8AXS1 | Q8axs1 anguilla ja |
| 27 | 45 | 41.7 | 488 | 16 Q8EYP2 | Q8eyp2 leptospira |
| 28 | 45 | 41.7 | 754 | 16 Q894A5 | Q894a5 clostridium |
| 29 | 45 | 41.7 | 880 | 16 Q8E6Y3 | Q8e6y3 streptococ |
| 30 | 45 | 41.7 | 880 | 16 Q8E1G6 | Q8e1g6 streptococ |
| 31 | 44 | 40.7 | 184 | 2 P94158 | P94158 synchococ |
| 32 | 44 | 40.7 | 264 | 10 Q8VWL5 | Q8vwl5 oryza sativ |
| 33 | 44 | 40.7 | 264 | 10 Q7XH02 | Q7xh02 oryza sativ |
| 34 | 44 | 40.7 | 414 | 16 Q8Y0E3 | Q8y0e3 raistonia s |
| 35 | 44 | 40.7 | 633 | 10 Q9SZX8 | Q9szx8 arabidopsis |
| 36 | 44 | 40.7 | 787 | 16 Q8G3V2 | Q8g3v2 bifidobacte |
| 37 | 44 | 40.7 | 878 | 16 Q8DVZ5 | Q8dvz5 streptococ |
| 38 | 44 | 40.7 | 880 | 16 Q9A1M8 | Q9a1m8 streptococ |
| 39 | 44 | 40.7 | 880 | 16 Q8P2S4 | Q8p2s4 streptococ |
| 40 | 44 | 40.7 | 880 | 16 Q8K8R3 | Q8k8r3 streptococ |
| 41 | 44 | 40.7 | 1099 | 16 Q9HYC2 | Q9hyc2 pseudomonas |
| 42 | 43.5 | 40.3 | 212 | 12 Q914H4 | Q914h4 sulfolobus |
| 43 | 43 | 39.8 | 93 | 6 Q862Y0 | Q862y0 bos taurus |
| 44 | 43 | 39.8 | 135 | 11 Q89059 | Q89059 mus musculu |
| 45 | 43 | 39.8 | 157 | 11 Q9CTR1 | Q9ctr1 mus musculu |

ALIGNMENTS

RESULT 1
Q86WL4
ID Q86WL4 PRELIMINARY; PRT; 518 AA.
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Gastric cancer multi-drug resistance protein variant.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Han Q., Wang X., Shi Y., Ding J., Fan D.;
RT "A variant of the 60-kD component of the human SS-A/Ro
ribonucleoprotein autoantigen involved in multidrug resistance of
gastric cancer cells";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY205314; AAO47001.1; -. 42C3508D263885A1 CRC64;
SQ SEQUENCE 518 AA; 58511 MW; 42C3508D263885A1 CRC64;

Query Match 100.0%; Score 108; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 6.1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KARIHPFHILIALETYKTH 20
Db 316 KARIHPFHILIALETYKTH 335

RESULT 2
Q86WL3
ID Q86WL3 PRELIMINARY; PRT; 534 AA.
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Gastric cancer multi-drug resistance protein variant.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Han Q., Wang X., Shi Y., Ding J., Fan D.;
 RT "A variant of the 60-kD component of the human SS-A/Ro
 RT ribonucleoprotein autoantigen involved in multidrug resistance of
 RT gastric cancer cells.";
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY205315; AA047002.1; -.
 SQ SEQUENCE 534 AA; 60197 MW; 63A4B96A77B6DD2D CRC64;

Query Match 100.0%; Score 108; DB 4; Length 534;
 Best Local Similarity 100.0%; Pred. No. 6.3e-10; Mismatches 0; Indels 0; Gaps 0;
 Matches 20; Conservative 0;

QY 1 KARIHPHILIALETYKTH 20
 |||||
 Db 316 KARIHPHILIALETYKTH 335

RESULT 3

Q9H1W6 PRELIMINARY; PRT; 538 AA.
 AC Q9H1W6;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE R101E13.2 (Sjogren syndrome antigen A2 (60kD, ribonucleoprotein
 DE autoantigen SS-A/Ro) (60kD, ribonucleoprotein autoantigen
 DE SS-A/Ro).
 GN SSA2.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wilson S.;
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;
 RC TISSUE-Testis;
 RA Strausberg R.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL136370; CAC17589.1; -.
 DR EMBL; BC036658; AAH36658.1; -.
 DR GO; GO:0030529; C:ribonucleoprotein complex; IEA.
 DR GO; GO:0019013; C:viral nucleocapsid; IEA.
 DR InterPro; IPR008858; TROVE.
 DR Pfam; PF05731; TROVE; 1.
 KW Nucleocapsid; Ribonucleoprotein.

SQ SEQUENCE 538 AA; 60670 MW; CD735B1DF2B13098 CRC64;
 Query Match 100.0%; Score 108; DB 4; Length 538;
 Best Local Similarity 100.0%; Pred. No. 6.3e-10; Mismatches 0; Indels 0; Gaps 0;
 Matches 20; Conservative 0;

QY 1 KARIHPHILIALETYKTH 20
 |||||
 Db 316 KARIHPHILIALETYKTH 335

RESULT 4

Q8R562 PRELIMINARY; PRT; 538 AA.
 AC Q8R562;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Ribonucleoprotein.

GN RO.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kannappell C.C., Fu S.M.;
 RT "Ribonucleoprotein Ro is similar to Ro/SS-A human complex.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; L81154; AAL77518.1; -.
 DR GO; GO:0019013; C:viral nucleocapsid; IEA.
 DR InterPro; IPR008858; TROVE.
 DR Pfam; PF05731; TROVE; 1.
 KW Nucleocapsid.

SQ SEQUENCE 538 AA; 60154 MW; 65E46CE43FA4572E CRC64;
 Query Match 91.7%; Score 99; DB 11; Length 538;
 Best Local Similarity 85.0%; Pred. No. 2.2e-08; Mismatches 1; Indels 0; Gaps 0;
 Matches 17; Conservative 2;

QY 1 KARIHPHILIALETYKTH 20
 |||||
 Db 316 KARIHPHILIALETYKTH 335

RESULT 5

Q80W12 PRELIMINARY; PRT; 538 AA.
 AC Q80W12;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Ssa2 protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RC MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klusner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.M., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millaby S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RC Strausberg R.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC051974; AAH51974.1; -.
 SQ SEQUENCE 538 AA; 60171 MW; 7850C6E41C726BDA CRC64;

Query Match 91.7%; Score 99; DB 11; Length 538;
 Best Local Similarity 85.0%; Pred. No. 2.2e-08; Mismatches 1; Indels 0; Gaps 0;
 Matches 17; Conservative 2;

QY 1 KARIHPHILIALETYKTH 20
 |||||
 Db 316 KARIHPHILIALETYKTH 335

```
[1]
RN      SEQUENCE FROM N.A.
RP     MEDLINE=22592660; PubMed=12705866;
RX     Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
RA     Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
RA     Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
RA     Ariakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA     Hatfull G.F.;
RT     "Origins of highly mosaic mycobacteriophage genomes.";
RL     Cell 113:171-182(2003).
DR     EMBL; AV129337; RAN16852.1; -.
SQ     SEQUENCE 552 AA; 60247 MW; B909A97F8159AAB7 CRC64;

Query Match          58.3%; Score 63; DB 9; Length 552;
Best Local Similarity 50.0%; Pred.No. 0.036;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY    1 KARHPFHILIALEYTKGH 20
      |||:|||::|||::|:
DB    314 KARVHPINVLAQRTYAOGY 333

RESULT 8
QBIP65 PRELIMINARY; PRT; 385 AA.
ID Q81P65
AC Q81P65;
DT 01-JUN-2003 (TrEMBLrel..24, Created)
DT 01-JUN-2003 (TrEMBLrel..24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel..25, Last annotation update)
DE Acetyltransferase, GNAT family.
GN BA2952.
OS Bacillus anthracis (strain Ames).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=198094;
RN [1]
SEQUENCE FROM N.A.
RP     MEDLINE=22608414; PubMed=12721629;
RX     Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,
RA     Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA     Holtzaple E.K., Okstad O.A., Helgason E., Ristone J., Wu M.,
RA     Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,
RA     DeBoy R.T., Madpu R., Dougherty S.C., Durkin A.S., Haft D.H.,
RA     Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA     Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA     Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA     Hazen A., Cline R., Redmond C., Thwaiter J.E., White O., Salzberg S.L.,
RA     Thomason B., Friedlander A.M., Koehler T.W., Hanna P.C., Kolsto A.-B.,
RA     Fraser C.M.;
RT     "The genome sequence of Bacillus anthracis Ames and comparison to
RL     closely related bacteria."
RL     Nature 423:81-86(2003).
DR     EMBL; AB017033; AAP26773.1; -.
DR     TIGR; BA2952; -.
DR     GO; GO:0008080; F:N-acetyltransferase activity; IEA.
DR     GO; GO:0016740; F:transferase activity; IEA.
DR     InterPro; IPR000182; GCN5Acetyl_trans.
DR     Pfam; PF00563; Acetyltransf_1.
KW Transferase; Complete proteome.
SQ SEQUENCE 385 AA; 45475 MW; AF1243CD8556FDF0 CRC64;

Query Match          44.4%; Score 48; DB 16; Length 385;
Best Local Similarity 50.0%; Pred.No. 9.4;
Matches 10; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

QY    2 ARH--PFHILIALEYTKG 19
      |||:|||:|||::|:
DB    55 AKHLIPPFIYIGKRFQMG 74

RESULT 9
COI858 PRELIMINARY; PRT; 997 AA.
ID Q01858
AC Q01858;
```


telomeres in an unusually compact eukaryotic genome, the cryptomonad nucleomorph.";
Proc. Natl. Acad. Sci. U.S.A. 97:200-205(2000).
[2]
SEQUENCE FROM N.A.
MEDLINE=21223349; PubMed=11323671;
Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L., Wu X., Reich M., Cavalier-Smith T., Maier U.;
"The highly reduced genome of an enslaved algal nucleus.";
Nature 410:1091-1096(2001).
EMBL; AJ010592; CAC27062.1; --
PIR; C90112; C9C112.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR001197; Ribosomal L10E.
DR Pfam; PF00826; Ribosomal L10e; 1.
DR TIGRfams; TIGR00279; L10e; 1.
DR Ribosomal protein.
KW SEQUENCE 188 AA; 21591 MW; 247D417F56289CF9 CRC64;
SQ
Query Match 43.5%; Score 47; DB 10; Length 188;
Best Local Similarity 70.0%; Pred. NO. 6;6;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 KARIHPPHIL 10
| | | | | | | | | |
DB 88 KIRIHPPHVI 97
| | | | | | | | | |
RESULT 11
Q8ILV2 PRELIMINARY; ERT; 219 AA.
ID Q8ILV2
AC Q8ILV2
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ribosomal protein L10, putative.
GN PF14.0141.
DE OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W., Carlton J.M., Paine A., Nelson K.E., Bowman S., Paulsen I.R., James K., Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S., Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S., Pertea M., Allen J., Selengut J.J., Haft D., Mather M.W., Vaidya A.B., Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A., McPadden G.I., Cummings L.M., Subramanian G.M., Mungall C., Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W., Praser C.M., Barrall B.;
"Genome sequence of the human malaria parasite Plasmodium falciparum";
Nature 419:498-511(2002).
RL EMBL; A5014818; AAN36753.1; --
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR001197; Ribosomal L10E.
DR Pfam; PF00826; Ribosomal L10e; 1.
DR TIGRfams; TIGR00279; L10e; 1.
DR PROSITE; PS01257; RIBOSOMAL L10E; 1.
DR SEQUENCE 219 AA; 25216 MW; 75C66FAD09FE3DDB CRC64;
SQ
Query Match 43.5%; Score 47; DB 5; Length 219;
Best Local Similarity 70.0%; Pred. NO. 7;8;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 KARIHPHIL 10
: |||||:
Db 88 RVRHPFVL 97

RESULT 12
ID O33728 PRELIMINARY; PRT; 266 AA.
AC O33728;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Dihydropteroate synthase (EC 2.5.1.15).
GN FOLP.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G72;
RX MEDLINE=98253959; PubMed=9593127;
RA Swedberg G., Ringertz S., Skoeld O.;
RT "Sulfonamide resistance in Streptococcus pyogenes is associated with
RT differences in the amino acid sequence of its chromosomal
RT dihydropteroate synthase.";
RL Antimicrob. Agents Chemother. 42:1062-1067(1998).
DR EMBL; AJ006807; CAA04243.1; -.
DR HSP; O05701; IAD1
DR GO; GO:0004156; F:dihydropteroate synthase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009396; P:folic acid and derivative biosynthesis; IEA.
DR InterPro; IPR00489; Dhdropt_synth.
DR HSP; O05701; IAD1
DR IPR; IPR006390; DHPS.
DR Pfam; PF00809; Pterin_bind.1.
DR TIGRFAMs; TIGR01496; DHPS_1.
DR PROSITE; PS00792; DHPS_1; 1.
DR PROSITE; PS00793; DHPS_2; 1.
DR KW Transferase.
SQ SEQUENCE 266 AA; 28659 MW; 9C953B16747C4763 CRC64;

Query Match 43.5%; Score 47; DB 2; Length 266;
Best Local Similarity 50.0%; Pred. No. 9.5;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KARIHPHILALETYKT 18
: |||||:
Db 80 KATIAKDVLSIDTYKT 97

RESULT 13
ID Q40650 PRELIMINARY; PRT; 173 AA.
AC Q40650;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Wilms' tumor-related protein QM (Fragment).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Indica variety Guang-lu-ai No.4; TISSUE=Seedling;
RX Zong H., Jiang Y., Cao K.;
RT "Isolation and Characterization of a rice QM gene, tumor suppressor or
RT development regulator.";
RL Thesis (1996), Biochemistry, Fudan University.
DR EMBL; U55212; AAA99158.1; -.
DR Gramene; Q40650; -.
```

```
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR001197; Ribosomal_L10E.
DR Pfam; PF00826; Ribosomal_L10e; 1.
DR PROSITE; PS01257; RIBOSOMAL_L10E; 1.
FT NON TER 1
SQ SEQUENCE 173 AA; 19262 MW; CEA0448DC6AE5DB9 CRC64;

Query Match 42.6%; Score 46; DB 10; Length 173;
Best Local Similarity 60.0%; Pred. No. 9;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPHIL 10
: |||||:
Db 37 RVRHPFVL 46

RESULT 14
ID Q7X9L9 PRELIMINARY; PRT; 177 AA.
AC Q7X9L9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE QM (Fragment).
OC Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. PH 82-2-2; TISSUE=Seed;
RX Li J.R., Wang F., Li Q.Z., Zhang X.S.;
RT "Gene isolation and expression of QM gene.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF470356; AAP80617.1; -.
FT NON TER 1
SQ SEQUENCE 177 AA; 19968 MW; B7E0A27711A0AE2B CRC64;

Query Match 42.6%; Score 46; DB 10; Length 177;
Best Local Similarity 60.0%; Pred. No. 9.2;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPHIL 10
: |||||:
Db 100 RVRHPFVL 109

RESULT 15
ID Q8MTT4 PRELIMINARY; PRT; 210 AA.
AC Q8MTT4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Ribosomal protein L10.
OS Entamoeba histolytica.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=5759;
RN [1]
RP SEQUENCE FROM N.A.
RX Chavez-Rios R., Arias-Romero L.E., Almaraz-Barrera M.,
RX Hernandez-Rivas R., Guillen N., Vargas M.;
RT "L10 Ribosomal Protein from Entamoeba histolytica Shares Structural
RT and Functional Homologies with QM/Jif-1: Proteins with Extraribosomal
RT Functions.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY072793; AAL68397.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005840; C:ribosome; IEA.
```

DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
 DR GO; GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro: IPR001197; Ribosomal_L10E.
 DR Pfam: PF00826; Ribosomal_L10e_1.
 DR TIGRFAMs: TIGR00279; L10e; 1.
 SQ SEQUENCE 210 AA; 23770 MW; 58117A6E39408819 CRC64;

Query Match 42.6%; Score 46; DB 5; Length 210;
 Best Local Similarity 70.0%; Pred. No. 11;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy . 1 KARIHPFHL 10
 : |||||:
 Db 88 RIRIHPFVL 97

Search completed: May 19, 2004, 15:46:25
 Job time : 43.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 15:28:52 ; Search time 53.625 seconds
(without alignments)
94.841 Million cell updates/sec

Title: US-10-005-684-7
Perfect score: 92
Sequence: 1 LSIALHVGDFHFEQLLSG 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 92 | 100.0 | 18 | 7 ADB81380 | Adb81380 Arthritis |
| 2 | 92 | 100.0 | 558 | 1 AAP70429 | Aap70429 Murine ne |
| 3 | 92 | 100.0 | 558 | 1 AAB36858 | Aab36858 Murine gl |
| 4 | 85 | 92.4 | 558 | 1 AAP70430 | Aap70430 Human neu |
| 5 | 85 | 92.4 | 558 | 2 AAY03633 | Aay03633 Hypoxia-r |
| 6 | 85 | 92.4 | 558 | 5 ABP65201 | Abp65201 Hypoxia-r |
| 7 | 85 | 92.4 | 558 | 6 ABR64229 | Abr64229 Angiogene |
| 8 | 85 | 92.4 | 558 | 6 ABR47477 | Abr47477 Breast ca |
| 9 | 85 | 92.4 | 558 | 6 ABU63752 | Abu63752 Human pro |
| 10 | 85 | 92.4 | 558 | 7 ADC69796 | Adc69796 Human pho |
| 11 | 85 | 92.4 | 558 | 7 ABW01152 | Abw01152 Human hyp |
| 12 | 79 | 85.9 | 548 | 6 ABUS0556 | Abu50556 Protein e |
| 13 | 71 | 77.2 | 549 | 6 ABM68987 | Abm68987 Photorhab |
| 14 | 70 | 76.1 | 551 | 6 AAO19955 | Aao19955 C elegans |
| 15 | 69 | 75.0 | 549 | 6 ABU39069 | Abu39069 Protein e |
| 16 | 68 | 73.9 | 548 | 6 ABU40692 | Abu40692 Protein e |
| 17 | 65 | 70.7 | 553 | 6 ABJ25440 | Abj25440 Aspergill |
| 18 | 63 | 68.5 | 553 | 6 ABJ26040 | Abj26040 Aspergill |
| 19 | 63 | 68.5 | 558 | 4 ABB58003 | Abb58003 Drosophil |
| 20 | 62 | 67.4 | 563 | 6 ABU30622 | Abu30622 Protein e |
| 21 | 62 | 67.4 | 592 | 5 ABB09167 | Abb09167 Methylo |
| 22 | 62 | 67.4 | 592 | 5 ABG61552 | Abg61552 High grow |
| 23 | 62 | 67.4 | 658 | 7 ADB70223 | Adb70223 C. neofor |
| 24 | 61 | 66.3 | 542 | 6 ABU28249 | Abu28249 Protein e |
| 25 | 61 | 66.3 | 546 | 6 ABU45435 | Abu45435 Protein e |

ALIGNMENTS

RESULT 1

ADB81380
ID ADB81380 standard; peptide; 18 AA.

XX ADB81380;

XX AC

XX DT 04-DEC-2003 (first entry)

XX DE Arthritis peptide used in a test for antibodies against arthritis.

XX KW immunoassay; autoimmune disease; autoantigen; ELISA;

XX KW enzyme-linked immunosorbant assay; saliva IgA; lupus; arthritis;

XX KW platelet glycoprotein; immune complex.

XX OS Synthetic.

XX US2003100035-A1.

XX PD 29-MAY-2003.

XX PF 08-NOV-2001; 2001US-00005684.

XX PR 08-NOV-2001; 2001US-00005684.

XX (VOJD/) VOJDANI A.

XX Vojdani A;

XX WPI; 2003-606630/57.

XX Saliva immunoassay for detection of antibodies for autoimmune disease, e.g. lupus, in patient, by determining level of antibodies against autoantigen, and comparing level of determined antibodies with normal levels of antibodies.

XX Example 8; Page 7; 16pp; English.

XX This invention relates to a novel saliva immunoassay for diagnosing an autoimmune disease in a patient. Specifically, the method determines the level of antibodies that are present against the autoantigens for a particular autoimmune disease and compares these results with normal levels to determine the likelihood or severity of such a disease. The test comprises a highly sensitive and accurate ELISA (enzyme-linked immunosorbant assay) that measures saliva IgA specific antibody titres against the purified antigens or a corresponding recombinant antigen or synthetic peptide of an autoantigen, where the autoantigens are lupus peptides, arthritis peptides, platelet glycoprotein or immune complexes. As such, this single test can be used to accurately detect antibodies for

| | | | | | | |
|----|------|------|-----|---|----------|-----------|
| 26 | 61 | 66.3 | 549 | 4 | AAU29348 | Novel mar |
| 27 | 61 | 66.3 | 549 | 6 | ABU47238 | Protein e |
| 28 | 61 | 66.3 | 549 | 6 | ABU47834 | Protein e |
| 29 | 61 | 66.3 | 549 | 6 | ABU31833 | Protein e |
| 30 | 61 | 66.3 | 549 | 6 | ABU15339 | Protein e |
| 31 | 60 | 65.2 | 554 | 6 | ABR52708 | Protein s |
| 32 | 54 | 58.7 | 550 | 6 | ABU49104 | Protein e |
| 33 | 53 | 57.6 | 550 | 5 | ABP73370 | Candida a |
| 34 | 51 | 55.4 | 556 | 6 | ABU17171 | Protein e |
| 35 | 51 | 55.4 | 565 | 6 | ADA33633 | Acinetoba |
| 36 | 50 | 54.3 | 554 | 6 | ABU41346 | Protein e |
| 37 | 49 | 53.3 | 554 | 6 | ABU38793 | Protein e |
| 38 | 46 | 50.0 | 554 | 6 | ABU40161 | Protein e |
| 39 | 46 | 50.0 | 554 | 6 | ABU40144 | Protein e |
| 40 | 45 | 48.9 | 298 | 6 | ADA34938 | Acinetoba |
| 41 | 45 | 48.9 | 744 | 2 | AAR52027 | Protein w |
| 42 | 44 | 47.8 | 556 | 6 | ABU19370 | Protein e |
| 43 | 43.5 | 47.3 | 323 | 6 | ABU22636 | Protein e |
| 44 | 43 | 46.7 | 103 | 6 | ADA34403 | Acinetoba |
| 45 | 43 | 46.7 | 540 | 6 | ABU22668 | Protein e |

CC diseases such as lupus or arthritis. Furthermore, it is useful to
 CC indicate ongoing pathology or to predict an early pathogenic reaction for
 CC autoimmune disease. This peptide sequence is the arthritis peptide used
 CC in a test for arthritis antibodies, in an exemplification of the
 CC invention.

XX SQ Sequence 18 AA;

Query Match 100.0%; Score 92; DB 7; Length 18;
 Best Local Similarity 100.0%; Pred. No. 2e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSIALHVGFDHFEQLLSG 18
 |||||
 Db 1 LSIALHVGFDHFEQLLSG 18
 |||||

RESULT 2
 AAP70429
 ID AAP70429 standard; protein; 558 AA.

XX AAP70429;

XX 25-MAR-2003 (revised)
 DT 16-JAN-1991 (first entry)

XX Murine neuroleukin.

DE Neuroleukin; neural cells; HIV; AIDS; env gene; vaccine.

XX Mus musculus.

XX Key Location/Qualifiers
 XX Region 1. 436
 FT /label= first exon
 FT Region 437
 FT /label= second exon

XX WO8707617-A.

XX 17-DEC-1987.

XX 04-JUN-1987; 87WO-US001323.

XX 05-JUN-1986; 86US-00872332.

XX (GEMV) GENETICS INST INC.

XX Gurney ME, Knopf JL;

XX WPI; 1987-362709/51.

XX N-PSDB; AAN70686.

XX New neuroleukin protein - useful for culturing neural cells, and for
 PT treating human immuno-deficiency virus.

XX Claim 1; Page 22; 29pp; English.

XX Neuroleukin is a single polypeptide chain of apparent mol. wt. 56000 +/-
 CC 2000 D (by SDS-PAGE), with the ability to maintain 1/2 maximal survival
 CC of spinal or sensory neurons cultured in vitro at a neuroleukin concn. of
 CC 0.000000125 M, and the ability to activate immunoglobulin secretion by
 CC peripheral blood lymphocytes. Neuroleukin is useful in tissue culture
 CC medium for culturing neural cells and extends the survival of sensory
 CC ganglia, brain cells and spinal neurons in culture. A portion of the DNA
 CC sequence has significant homology to portion of the HTLV III/LAV envelope
 CC protein gene, thus may be useful to treat AIDS, or in an AIDS vaccine.
 CC Its treatment of HIV infection is claimed. The protein is encoded by 2
 CC exons. (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 558 AA;

Query Match 100.0%; Score 92; DB 1; Length 558;

Best Local Similarity 100.0%; Pred. No. 8.8e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSIALHVGFDHFEQLLSG 18
 |||||
 Db 282 LSIALHVGFDHFEQLLSG 299
 |||||

RESULT 3
 AAB36858
 ID AAB36858 standard; protein; 558 AA.

XX AAB36858;

XX 20-FEB-2001 (first entry)

XX Murine glucose-6-phosphate isomerase.

XX Glucose-6-phosphate isomerase; GPI; arthritis; mouse.

XX Mus sp.

XX WO200064469-A2.

XX 02-NOV-2000.

XX 21-APR-2000; 2000WO-IB000600.

XX 22-APR-1999; 99US-0130730P.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX Benoit C, Matsumoto I, Korganow A, Mathis D, Maccioni M, Ji H;

XX WPI; 2000-679643/66.

XX Detecting the presence of autoantibodies against glucose-6-phosphate
 PT isomerase or GPI-like protein in a patient's sample, for diagnosing
 PT arthritis.

XX Disclosure; Fig 4; 37pp; English.

XX The present invention relates to diagnosing arthritis. The method
 CC involves detecting the presence of autoantibodies against glucose-6-
 CC phosphate isomerase (GPI) or GPI-like proteins, in the plasma or serum
 CC sample of the patient. Antibodies may also be developed for the
 CC diagnosis, treatment and prevention of arthritis

XX SQ Sequence 558 AA;

Query Match 100.0%; Score 92; DB 3; Length 558;
 Best Local Similarity 100.0%; Pred. No. 8.8e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LSIALHVGFDHFEQLLSG 18
 |||||
 Db 282 LSIALHVGFDHFEQLLSG 299
 |||||

RESULT 4
 AAP70430
 ID AAP70430 standard; protein; 558 AA.

XX AAP70430;

XX 25-MAR-2003 (revised)

DT 16-JAN-1991 (first entry)

XX Human neuroleukin.

XX Neuroleukin; neural cells; HIV; AIDS; env gene; vaccine.

XX Homo sapiens.

XX FH Key Location/Qualifiers
 XX PA 1. .337
 XX FT /label= first exon
 XX FT 338. .558
 XX FT /label= second exon
 XX XX WO8707617-A.
 XX XX
 XX PD 17-DEC-1987.
 XX XX
 XX PF 04-JUN-1987; 87WO-US001323.
 XX XX
 XX PR 05-JUN-1986; 86US-00872332.
 XX XX
 XX PA (GEMY) GENETICS INST INC.
 XX XX
 XX PI Gurney ME, Knopf JL;
 XX XX
 XX DR WPI; 1987-362709/51.
 XX DR N-PSDB; AAN70687.
 XX XX
 XX PT New neuro:leukin protein - useful for culturing neural cells, and for
 XX PT treating human immuno-deficiency virus.
 XX XX
 XX PS Claim 1; Page 22; 29pp; English.
 XX XX
 XX CC Neuroleukin is a single polypeptide chain of apparent mol. wt. 56000 +/-
 XX CC 2000 D (by SDS-PAGE), and with ability to maintain 1/2 maximal survival
 XX CC of spinal or sensory neurons cultured in vitro at a neuroleukin concn. of
 XX CC 0.0000000125 M, and the ability to activate immunoglobulin secretion by
 XX CC peripheral blood lymphocytes. Neuroleukin is useful in tissue culture
 XX CC medium for culturing neural cells and extends the survival of sensory
 XX CC ganglia, brain cells and spinal neurons in culture. A portion of the DNA
 XX CC sequence has significant homology to portion of the HTLV III/LAV envelope
 XX CC protein gene, thus may be useful to treat AIDS, or in an AIDS vaccine.
 XX CC Its treatment of HIV infection is claimed. The protein is expressed from
 XX CC 2 exons. (Updated on 25-MAR-2003 to correct PA field.)
 XX XX
 XX SQ Sequence 558 AA;
 XX
 XX Query Match 92.4%; Score 85; DB 1; Length 558;
 XX Best Local Similarity 94.4%; Pred. No. 1.6e-06;
 XX Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 LSIALHVGDFHFEQLLSG 18
 XX |||||
 XX Db 282 LSIALHVGDFNFEQLLSG 299
 XX
 XX RESULT 5
 XX AAY03633
 XX ID AAY03633 standard; protein; 558 AA.
 XX AC AAY03633;
 XX XX
 XX DT 04-JUN-1999 (first entry)
 XX XX
 XX DE Hypoxia-regulated gene RFP971 product.
 XX XX
 XX KW Hypoxia-regulated gene; therapeutic; diagnostic; hypoxia; ischemia;
 XX KW apoptosis; angiogenesis; tumorigenic cell; trauma; limb reattachment;
 XX KW revascularisation.
 XX XX
 XX OS Homo sapiens.
 XX XX
 XX PN WO9909049-A1.
 XX XX
 XX PD 25-FEB-1999.
 XX XX
 XX PF 21-AUG-1998; 98WO-US017296.
 XX XX
 XX PR 21-AUG-1997; 97US-0056453P.

XX XX (QUAR-) QUARK BIOTECH INC.
 XX PA (KOHN/) KOHN K I.
 XX XX
 XX PI Binat P, Skaliter R;
 XX XX
 XX DR WPI; 1999-180965/15.
 XX DR N-PSDB; AAX29142.
 XX XX
 XX PT New isolated hypoxia-related genes - used to develop products for use in
 XX PT therapy and diagnosis in e.g. hypoxia, ischaemia, apoptosis and
 XX PT angiogenesis.
 XX XX
 XX PS Claim 14; Page 80-81; 92pp; English.
 XX XX
 XX CC Sequences AAY03632-36 represent products of hypoxia-regulated genes of
 XX CC the invention. The genes and their products can be used therapeutically
 XX CC and diagnostically in hypoxia, ischemia, apoptosis and angiogenesis. The
 XX CC products and methods can be used for e.g. inducing apoptosis in
 XX CC tumorigenic cells or angiogenesis in trauma situations where e.g. a limb
 XX CC must be reattached or in a transplant where revascularisation is needed
 XX XX
 XX SQ Sequence 558 AA;
 XX
 XX Query Match 92.4%; Score 85; DB 2; Length 558;
 XX Best Local Similarity 94.4%; Pred. No. 1.6e-06;
 XX Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 LSIALHVGDFHFEQLLSG 18
 XX |||||
 XX Db 282 LSIALHVGDFNFEQLLSG 299
 XX
 XX RESULT 6
 XX ABP65201
 XX ID ABP65201 standard; protein; 558 AA.
 XX AC ABP65201;
 XX XX
 XX DT 12-NOV-2002 (first entry)
 XX XX
 XX DE Hypoxia-regulated protein #75.
 XX XX
 XX KW Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;
 XX KW antiinflammatory; vulnery; gynecological; ophthalmological; vaccine;
 XX KW hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;
 XX KW ischaemic condition; reperfusion injury; retinopathy; neonatal stress;
 XX KW preclapmsia; atherosclerosis; inflammatory condition; wound healing;
 XX KW inflammation; erythropoiesis; hair loss; human.
 XX XX
 XX OS Homo sapiens.
 XX XX
 XX PN WO200246465-A2.
 XX XX
 XX PD 13-JUN-2002.
 XX XX
 XX PF 10-DEC-2001; 2001WO-GB005458.
 XX XX
 XX PR 08-DEC-2000; 2000GB-00030076.
 XX PR 08-FEB-2001; 2001GB-00003156.
 XX PR 25-OCT-2001; 2001GB-00025666.
 XX XX
 XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX XX
 XX PI White J, Mundy CR, Ward NR, Krige D, Kingsman SM, Harris RA;
 XX PI Rayner WN;
 XX XX
 XX DR WPI; 2002-627238/67.
 XX XX
 XX PT Identifying a gene involved in disease for treating hypoxia-regulated
 XX PT conditions, comprises comparing the transcriptome/proteome of two cell
 XX PT types under different conditions and identifying a differentially
 XX PT regulated gene.

PS Claim 35; Page 390; 538pp; English.

XX The present invention relates to methods for identifying genes and

XX proteins that are implicated in a specific disease or physiological

CC condition. The method comprises comparing the transcriptome/proteome of a

CC specialised cell type implicated in a disease or condition with that of a

CC second specialised cell type, under two experimental conditions, and

CC identifying a gene that is differentially regulated in the two

CC specialised cell types under experimental conditions. ABV77873-ABV78116

CC and ABP65061-ABP65257 were identified using the methods of the invention.

CC The coding sequences and proteins are useful for treating a disease in a

CC patient, for manufacture of a medicament for treating hypoxia-regulated

CC conditions, and for regulating tumourigenesis, angiogenesis, apoptosis,

CC biological response to hypoxia conditions, or hypoxic-associated

CC pathology in a patient. The coding sequences and proteins are also useful

CC for monitoring the therapeutic treatment of a disease or physiological

CC condition, such as cancer, ischaemic conditions, reperfusion injury,

CC retinopathy, neonatal stress, preclampsia, atherosclerosis, inflammatory

CC conditions, wound healing, inflammation, erythropoiesis or hair loss

XX

SQ Sequence 558 AA;

Query Match 92.4%; Score 85; DB 5; Length 558;

Best Local Similarity 94.4%; Pred. No. 1.6e-06;

Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHFEQLLSG 18

Db 282 LSIALHVGDFHFEQLLSG 299

RESULT 7

ABR64229

ID ABR64229 standard; protein; 558 AA.

XX ABR64229;

XX 15-OCT-2003 (first entry)

XX Angiogenesis protein BNO134.

XX Cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;

XX antipsoriatic; antiarteriosclerotic; cardiast; vasotropic; angiogenesis;

XX gene therapy; vasculature; cancer; rheumatoid arthritis; psoriasis;

XX diabetic retinopathy; cardiovascular disease; atherosclerosis;

XX ischemic limb disease; coronary artery disease.

XX Homo sapiens.

XX WO2003027285-A1.

XX 03-APR-2003.

XX 19-SEP-2002; 2002WO-AU001282.

XX 27-SEP-2001; 2001AU-00007973.

XX 27-SEP-2001; 2001AU-00007974.

XX 11-OCT-2001; 2001AU-00008210.

XX 29-OCT-2001; 2001AU-00008532.

XX 13-NOV-2001; 2001AU-00008838.

XX 28-AUG-2002; 2002AU-00951032.

XX (BION-) BIONOMICS LTD.

XX Gamble JR, Hahn CN, Vadas MA;

XX WPI; 2003-354655/33.

XX N-PSDB; ACF34504.

XX New angiogenic genes and polypeptides, useful for diagnosing,

PT prognosticating or treating an angiogenesis-related disorder, e.g.

PT cancer, rheumatoid arthritis, diabetic retinopathy, psoriasis or

PT cardiovascular diseases.

PS Claim 15; SEQ ID NO 164; 90pp; English.

XX The invention relates to the isolation of novel genes (ACF34446-ACF34559)

XX encoding proteins (ABR64180-ABR64281) involved in the process of

CC angiogenesis. The nucleic acid molecules are useful in identifying and/or

CC obtaining full-length human genes involved in an angiogenic process. The

CC nucleic acid molecule, polypeptides or complexes encoded, cells or

CC genetically modified non-human animals derived from these are useful for

CC the screening of candidate pharmaceutical compounds used in treating

CC angiogenesis-related disorders. They are also useful for diagnosing,

CC prognosticating or treating an angiogenesis-related disorder, which

CC involves uncontrolled or enhanced angiogenesis or is a disorder in which

CC a decreased vasculature is of benefit (e.g. cancer, rheumatoid arthritis,

CC diabetic retinopathy, psoriasis or cardiovascular diseases such as

CC atherosclerosis), or involves inappropriately arrested or decreased

CC angiogenesis or is a disorder in which an expanding vasculature is of

CC benefit (e.g. ischemic limb disease or coronary artery disease). The

CC modulator of expression or activity of the polypeptide encoded by the

CC nucleic acid sequence is useful for manufacturing a medicament for the

CC treatment of an angiogenesis-related disorder. This sequence corresponds

CC to one of the novel angiogenic protein

SQ Sequence 558 AA;

Query Match 92.4%; Score 85; DB 6; Length 558;

Best Local Similarity 94.4%; Pred. No. 1.6e-06;

Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LSIALHVGDFHFEQLLSG 18

Db 282 LSIALHVGDFHFEQLLSG 299

RESULT 8

ABR47477

ID ABR47477 standard; protein; 558 AA.

XX ABR47477;

XX 12-JUN-2003 (first entry)

XX Breast cancer associated protein sequence SEQ ID NO:188.

XX Human; breast cancer; cytostatic; gene therapy.

XX Homo sapiens.

XX WO2003004989-A2.

XX 16-JAN-2003.

XX 21-JUN-2002; 2002WO-US019669.

XX 21-JUN-2001; 2001US-0299887P.

XX 27-JUN-2001; 2001US-0301572P.

XX 18-JUL-2001; 2001US-0306501P.

XX 25-SEP-2001; 2001US-0325002P.

XX 05-MAR-2002; 2002US-0362585P.

XX 14-MAY-2002; 2002US-0380391P.

XX (MILL-) MILLENIUM PHARM INC.

XX Lillie J, Gannavarapu M, Glatt K, Hoersch S, Kamatkar S;

XX Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;

XX Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;

XX WPI; 2003-210381/20.

XX N-PSDB; ACC50171.

XX Breast cancer diagnosis or treatment by comparing the level of expression

PT of a marker in a patient sample with that in the control non-breast

PT cancer sample.
XX Claim 1; SEQ ID NO 188; 128pp; English.
XX
XX The present invention describes a method for assessing whether a patient
XX is afflicted with breast cancer. The method comprises comparing the level
XX of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
XX ABR47386 to ABR47632) in a patient sample and the normal level of
XX expression of the marker in a control non-breast cancer sample, where a
XX significant increase in the level of expression of the marker in the
XX patient sample and the normal level is an indication that the patient is
XX afflicted with breast cancer. The breast cancer associated sequences from
XX the present invention have cytostatic activities and can be used in gene
XX therapy. The method is useful for diagnosing and treating breast cancer.
XX N.B. The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 558 AA;
XX
XX Query Match 92.4%; Score 85; DB 6; Length 558;
XX Best Local Similarity 94.4%; Pred. No. 1.6e-06;
XX Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 LSTALHVGFDHFEQLLSG 18
XX |||||:|||||
XX Db 282 LSTALHVGFDHFEQLLSG 299
XX
XX RESULT 9
XX ABU63752
XX ID ABU63752 standard; protein; 558 AA.
XX AC ABU63752;
XX
XX DT 15-OCT-2003 (first entry)
XX
XX DE Human protein from gene upregulated in hypoxia, RTP971.
XX
XX KW Human; hypoxia; RTP971; cerebroprotective; cardiatic; cytostatic;
XX KW ischaemia; stroke; apoptotic disease; tumour; angiogenesis; neuroleukin.
XX
XX OS Homo sapiens.
XX
XX PN US2003104973-A1.
XX
XX PD 05-JUN-2003.
XX
XX PF 06-MAR-2002; 2002US-00091333.
XX
XX PR 21-AUG-1997; 97US-0056453P.
XX PR 21-AUG-1998; 98US-00138112.
XX PR 28-JUN-2000; 2000US-00604978.
XX
XX PA (QUAR-) QUARK BIOTECH INC.
XX
XX PI Einat P, Skaliter R;
XX
XX DR WPI; 2003-576968/54.
XX DR N-PSDB; ACH00175.
XX
XX PT New hypoxia-regulated nucleic acid molecules and polypeptides, useful for
XX PT treating hypoxia or ischemia-related disease, such as stroke.
XX
XX PS Example 1; Page 30-32; 50pp; English.
XX
XX CC The invention relates to an isolated hypoxia regulated nucleic acid
XX CC molecule comprising a rat or human sequence (designated RTP801 and
XX CC RTP779) encoding a polypeptide appearing as ABU63753 and ABU63754. Also
XX CC included are an antibody that specifically binds to the polypeptide
XX CC encoded by the nucleic acid molecule and treating hypoxia or ischaemia-
XX CC related disease in a subject by administering to the subject an
XX CC antagonist of a protein comprising of the human protein or its analogue.

CC The nucleic acid molecule, and protein or its antagonist are useful for
CC treating hypoxia or ischaemia-related disease, e.g. stroke, apoptotic
CC disease (e.g. tumours) and diseases involving angiogenesis. The present
CC sequence is encoded by a disclosed novel human hypoxia regulated cDNA
CC RTP971, thought to be the Neuroleukin gene
XX
XX SQ Sequence 558 AA;
XX
XX Query Match 92.4%; Score 85; DB 6; Length 558;
XX Best Local Similarity 94.4%; Pred. No. 1.6e-06;
XX Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 LSTALHVGFDHFEQLLSG 18
XX |||||:|||||
XX Db 282 LSTALHVGFDHFEQLLSG 299
XX
XX RESULT 10
XX ADC69796
XX ID ADC69796 standard; protein; 558 AA.
XX AC ADC69796;
XX
XX DT 18-DEC-2003 (first entry)
XX
XX DE Human phosphohexose isomerase enzyme encoded by RTP971 cDNA clone.
XX
XX KW human; RTP971; hypoxia regulated; neuroleukin; astrocyte motility factor;
XX KW gene therapy; ischaemia; angiogenesis; apoptosis; cancer; angiogenic;
XX KW apoptotic; vasodilator; antisense; phosphohexose isomerase; enzyme.
XX
XX OS Homo sapiens.
XX
XX PN US6555667-B1.
XX
XX PD 29-APR-2003.
XX
XX PF 28-JUN-2000; 2000US-00604728.
XX
XX PR 21-AUG-1997; 97US-0056453P.
XX PR 21-AUG-1998; 98US-00138112.
XX
XX PA (QUAR-) QUARK BIOTECH INC.
XX
XX PI Binat P, Skaliter R;
XX
XX DR WPI; 2003-786295/74.
XX DR N-PSDB; ADC69805.
XX
XX PT New isolated hypoxia-responding nucleic acid for use in regulating
XX PT angiogenesis or apoptosis, or for diagnosing or treating ischemia in a
XX PT patient.
XX
XX PS Example; SEQ ID NO 8; 33pp; English.
XX
XX CC This invention relates to novel hypoxia regulated genes. Specifically, it
XX CC refers to neuroleukin, a hypoxia regulated gene that is also an astrocyte
XX CC motility factor important in the formation of blood vessels in the brain
XX CC and retina, accordingly antisense inhibitors can be useful for the
XX CC diagnosis and treatment of angiogenesis. The present invention describes
XX CC the identification of candidate genes and gene products that can through
XX CC gene therapy be used for the treatment of hypoxia and ischaemia, and
XX CC furthermore can regulate apoptosis, and therefore also be used to treat
XX CC cancer. Accordingly, they have various activities such as angiogenic and
XX CC apoptotic, as well as being vasodilators. This polypeptide sequence is
XX CC the human phosphohexose isomerase enzyme, a neuroleukin protein homologue
XX CC encoded by the RTP971 cDNA clone of the invention.
XX
XX SQ Sequence 558 AA;
XX
XX Query Match 92.4%; Score 85; DB 7; Length 558;
XX Best Local Similarity 94.4%; Pred. No. 1.6e-06;
XX Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY      1  LSTALHVGFDHFEQLLSG 18
DB      282 LSTALHVGFDNFQQLSG 299
          |||||:|||||
          |||||:|||||

RESULT 11
ABW01152
ID  ABW01152 standard; protein; 558 AA.
XX  AC
XX  ABW01152;
XX  DT 15-JAN-2004 (first entry)
XX  Human hypoxia-regulating protein from clone RTP971.
DE
XX  Hypoxia-regulating gene; hypoxia; ischaemia; angiogenesis; apoptosis;
XX  cytosstatic; vasotropic; gene therapy; human.
XX  OS Homo sapiens.
XX  US2003124116-A1.
XX  PN
XX  PD 03-JUL-2003.
XX  PF
XX  PP 23-DEC-2002; 2002US-00325878.
XX  PR 21-AUG-1997; 97US-0056453P.
XX  PR 21-AUG-1998; 98US-00138112.
XX  PR 28-JUN-2000; 2000US-00604728.
XX  (QUAR-) QUARK BIOTECH INC.
XX  PA
XX  PI
XX  PI Einat P, Skaliter R;
XX  DR WPI; 2003-810969/76.
XX  DR N-PSDB; AAD62298.
XX  XX
XX  New hypoxia-regulating genes and proteins, useful for diagnosing and
PT  treating hypoxia and ischemia, for regulating angiogenesis or apoptosis,
PT  or for regulating response to hypoxic conditions in a patient.
PT  Example; Page 22-24; 34pp; English.
XX  PS
XX  PS The present invention provides new hypoxia-regulating genes and proteins.
CC  The invention is useful for diagnosing, treating hypoxia and ischaemia,
CC  for regulating angiogenesis or apoptosis and for regulating response to
CC  hypoxic conditions in a patient. The invention is also useful in gene
CC  therapy. The present sequence is human hypoxia-regulating protein from
CC  clone RTP971
XX  SQ
XX  SQ Sequence 558 AA;
XX
XX  Query Match 92.4%; Score 85; DB 7; Length 558;
XX  Best Local Similarity 94.4%; Pred. No. 1.6e-06;
XX  Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1  LSTALHVGFDHFEQLLSG 18
DB      282 LSTALHVGFDNFQQLSG 299
          |||||:|||||
          |||||:|||||

RESULT 12
ABU05056
ID  ABU05056 standard; protein; 548 AA.
XX  AC
XX  ABU05056;
XX  DT 19-JUN-2003 (first entry)
XX  Protein encoded by Prokaryotic essential gene #36083.
XX  DE
XX  XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX  XX

```

```

ABM68987
ID  ABM68987 standard; protein; 549 AA.
XX  AC
XX  AC
XX  ABM68987;
XX
XX  20-NOV-2003 (first entry)
XX  DE
XX  DE
XX  Photorhabdus luminescens protein sequence #2084.
XX  DE
XX  Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
XX  KW detection; food; gene expression; plant; animal; microorganism; toxin;
XX  KW antibiotic; biopesticide; virulence factor; disease model; plague;
XX  KW whooping cough.
XX
XX  Photorhabdus luminescens.
XX  OS
XX  PN WO200294867-A2.
XX  XX
XX  28-NOV-2002.
XX  PD
XX  XX
XX  07-FEB-2002; 2002WO-1B003040.
XX  PF
XX  XX
XX  07-FEB-2001; 2001FR-00001659.
XX  PR
XX  XX
XX  (INSP ) INST PASTEUR.
XX  PA
XX  (CNRS ) CNRS CENT NAT RECH SCI.
XX  PA
XX  Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A,
XX  PI Buchrieser C;
XX  PI
XX  WPI; 2003-148459/14.
XX  DR
XX  XX
XX  Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
XX  PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
XX  PT
XX  Claim 2; SEQ ID NO 2084; 1205pp; French.
XX  PS
XX  XX
XX  The invention relates to the isolation of genes and their encoded
XX  CC proteins from Photorhabdus luminescens. The isolated sequences are
XX  CC sources of probes and primers for detecting the genome of P. luminescens
XX  CC and related species; to study polymorphisms; for gene analysis and for
XX  CC detection/amplification of the genes. Antibodies (Ab) raised against the
XX  CC polypeptides encoded by the genes are used for detection/identification
XX  CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
XX  CC carry a gene-containing vector are used to select compounds that
XX  CC modulate, regulate, induce or inhibit expression of the genes in plants.
XX  CC animals or microorganisms other than P. luminescens and are able to alter
XX  CC response or sensitivity to toxins and antibiotics produced by P.
XX  CC luminescens. Cells transformed to express the genes are useful for
XX  CC recombinant production of the proteins, particularly toxins and
XX  CC antibacterials useful as insecticides, bactericides and fungicides. The
XX  CC genes, proteins, vectors containing the genes and Ab are also useful
XX  CC therapeutically (to treat microbial infection by bacteria or fungi that
XX  CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
XX  CC biopesticides. Other uses of the genes and the proteins are as virulence
XX  CC factors and for identifying targets of human diseases for which P.
XX  CC luminescens is a model (particularly plague and whooping cough). This
XX  CC sequence represents one of the isolated P. luminescens proteins
XX  CC
XX  Sequence 549 AA;
XX  SQ
XX  Query Match 77.2%; Score 71; DB 6; Length 549;
XX  Best Local Similarity 77.8%; Pred. No. 0.00051;
XX  Matches 14; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
XX
XX  QY 1 LSIALHVGDFHPEQLLSG 18
XX  |||||:|||||:|||||
XX  Db 279 LSIALSIGFENFEQLLSG 296
XX
XX  RESULT 14
XX  AAO19955
XX  ID AAO19955 standard; protein; 551 AA.
XX
AAO19955;
XX
XX  11-AUG-2003 (first entry)
XX  DE
XX  DE
XX  C elegans phosphoglucose isomerase protein.
XX  XX
XX  Adult lifespan; glycolysis; mitochondrial respiratory chain; aging;
XX  KW gene therapy; cytochrome C1; NADH oxidoreductase; ATP synthase; aging;
XX  KW cytochrome C oxidase; phosphoglucose isomerase; GTPase; LLMW-1; LLMW-2;
XX  LLMW-3; LLMW-4; HSF-1.
XX  XX
XX  Caenorhabditis elegans.
XX  OS
XX  PN WO2003000861-A2.
XX  XX
XX  03-JAN-2003.
XX  PD
XX  XX
XX  24-JUN-2002; 2002WO-US020247.
XX  PF
XX  XX
XX  22-JUN-2001; 2001US-0300577P.
XX  PR
XX  25-JUN-2001; 2001US-0301052P.
XX  PR
XX  18-APR-2002; 2002US-0373975P.
XX  PR
XX  XX
XX  (REGC ) UNIV CALIFORNIA.
XX  PA
XX  Kenyon C, Apfeld J, Dillin A, Garigan D, Hsu AA,
XX  PI Lehrer-Graiwer J, Murphy C;
XX  PI
XX  WPI; 2003-229334/22.
XX  DR
XX  XX
XX  Identifying a compound that modulates aging, comprises contacting the
XX  PT compound with a polypeptide comprising, e.g. NADH oxidoreductase or
XX  PT phosphoglucose isomerase, and determining the effect of the compound on
XX  PT the polypeptide.
XX  PT
XX  Claim 1; Page 88; 112pp; English.
XX  PS
XX  XX
XX  The present invention relates to a method of identifying a compound that
XX  CC modulates aging, involving contacting a compound with one of cytochrome
XX  CC C1, NADH oxidoreductase, ATP synthase, cytochrome C oxidase,
XX  CC phosphoglucose isomerase, GTPase, LLMW-1, LLMW-2, LLMW-3, LLMW-4 and HSF-1,
XX  CC and determining the functional effect upon the polypeptide. The method is
XX  CC useful identifying compounds that modulate aging and in regulating adult
XX  CC lifespan in eukaryotes, e.g. humans, animals and plants. The compound may
XX  CC also be used in diagnosis and therapy related to lifespan extension, life
XX  CC expectancy, and aging. The present sequence is the C. elegans
XX  CC phosphoglucose isomerase, which is one of the proteins involved in
XX  CC lifespan regulation and described in the invention
XX  CC
XX  Sequence 551 AA;
XX  SQ
XX  Query Match 76.1%; Score 70; DB 6; Length 551;
XX  Best Local Similarity 66.7%; Pred. No. 0.00078;
XX  Matches 12; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
XX
XX  QY 1 LSIALHVGDFHPEQLLSG 18
XX  |||||:|||||:|||||
XX  Db 280 LSIALHVGDFNVEKLLDG 297
XX
XX  RESULT 15
XX  ABU39069
XX  ID ABU39069 standard; protein; 549 AA.
XX  XX
XX  AC ABU39069;
XX  XX
XX  19-JUN-2003 (first entry)
XX  DT
XX  XX
XX  Protein encoded by Prokaryotic essential gene #24596.
XX  DE
XX  XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX  XX

```

OS Pasteurella multocida.
 XX WO200277183-A2.
 XX
 XX 03-OCT-2002.
 XX
 XX 21-MAR-2002; 2002WO-US009107.
 XX
 XX 21-MAR-2001; 2001US-00815242.
 XX 06-SEP-2001; 2001US-00948993.
 XX 25-OCT-2001; 2001US-0342923P.
 XX 08-FEB-2002; 2002US-00072851.
 XX 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 XX
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 XX N-PSDB; ACA42939.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 XX for homologous nucleic acids required for cellular proliferation to
 XX isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 25; SEQ ID NO 66993; 1766pp; English.
 XX
 XX The invention relates to an isolated nucleic acid comprising any one of
 XX the 6213 antisense sequences given in the specification where expression
 XX of the nucleic acid inhibits proliferation of a cell. Also included are:
 XX (1) a vector comprising a promoter operably linked to the nucleic acid
 XX encoding a polypeptide whose expression is inhibited by the antisense
 XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
 XX polypeptide or its fragment whose expression is inhibited by the
 XX antisense nucleic acid; (4) an antibody capable of specifically binding
 XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 XX proliferation or the activity of a gene in an operon required for
 XX proliferation; (7) identifying a compound that influences the activity of
 XX the gene product or that has an activity against a biological pathway
 XX required for proliferation, or that inhibits cellular proliferation; (8)
 XX identifying a gene required for cellular proliferation or the biological
 XX pathway in which a proliferation-required gene or its gene product lies
 XX or a gene on which the test compound that inhibits proliferation of an
 XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
 XX compound's activity; (11) a culture comprising strains in which the gene
 XX product is overexpressed or underexpressed; (12) determining the extent
 XX to which each of the strains is present in a culture or collection of
 XX strains; or (13) identifying the target of a compound that inhibits the
 XX proliferation of an organism. The antisense nucleic acids are useful for
 XX identifying proteins or screening for homologous nucleic acids required
 XX for cellular proliferation to isolate candidate molecules for rational
 XX drug discovery programs, or for screening homologous nucleic acids
 XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 XX the target prokaryotic essential genes. Note: The sequence data for this
 XX patent did not form part of the printed specification, but was obtained
 XX in electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 549 AA;
 XX
 XX Query Match 75.0%; Score 69; DB 6; Length 549;
 XX Best Local Similarity 72.2%; Pred. No. 0.0012;
 XX Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 XX
 XX Qy 1 LSIALHVGFDHFQQLSG 18
 XX Db 279 LSIALSIGFHFALLAG 296
 XX
 XX Search completed: May 19, 2004, 15:43:58
 XX Job time : 55.625 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 15:29:47 ; Search time 9.58333 Seconds
(without alignments)
108.668 Million cell updates/sec

Title: US-10-005-684-6

Perfect score: 108
Sequence: 1 KARIHFPHILIAETVKTGH 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 108 | 100.0 | 538 | 1 | P10155 homo sapien |
| 2 | 101 | 93.5 | 538 | 1 | P42700 xenopus lae |
| 3 | 99 | 91.7 | 538 | 1 | O0848 mus musculus |
| 4 | 46 | 42.6 | 150 | 1 | Q40592 nicotiana t |
| 5 | 46 | 42.6 | 218 | 1 | R102 ORISA |
| 6 | 46 | 42.6 | 219 | 1 | R103 ORISA |
| 7 | 46 | 42.6 | 219 | 1 | R103 ORISA |
| 8 | 46 | 42.6 | 219 | 1 | R103 ORISA |
| 9 | 46 | 42.6 | 220 | 1 | R103 ORISA |
| 10 | 46 | 42.6 | 220 | 1 | R103 ORISA |
| 11 | 46 | 42.6 | 221 | 1 | R103 ORISA |
| 12 | 46 | 42.6 | 228 | 1 | R103 ORISA |
| 13 | 46 | 42.6 | 877 | 1 | DPO1 LACLA |
| 14 | 46 | 42.6 | 877 | 1 | DPO1 LACLA |
| 15 | 45 | 41.7 | 221 | 1 | R103 ORISA |
| 16 | 45 | 41.7 | 252 | 1 | R103 ORISA |
| 17 | 45 | 41.7 | 643 | 1 | R060 CAEL |
| 18 | 45 | 41.7 | 877 | 1 | DPO1 STRP |
| 19 | 45 | 41.7 | 877 | 1 | DPO1 STRP |
| 20 | 44 | 40.7 | 304 | 1 | ERA BACHD |
| 21 | 44 | 40.7 | 2567 | 1 | M188 HUMAN |
| 22 | 43 | 39.8 | 210 | 1 | R103 ORISA |
| 23 | 43 | 39.8 | 213 | 1 | R103 ORISA |
| 24 | 43 | 39.8 | 213 | 1 | R103 ORISA |
| 25 | 43 | 39.8 | 213 | 1 | R103 ORISA |
| 26 | 43 | 39.8 | 371 | 1 | H2AW HUMAN |
| 27 | 43 | 39.8 | 759 | 1 | C102 MOUSE |
| 28 | 43 | 39.8 | 852 | 1 | C102 MOUSE |
| 29 | 43 | 39.8 | 872 | 1 | C102 HUMAN |
| 30 | 43 | 39.8 | 1150 | 1 | AN61 MOUSE |
| 31 | 43 | 39.8 | 1331 | 1 | XDH BOVIN |
| 32 | 42.5 | 39.4 | 791 | 1 | VD05 FOWPV |
| 33 | 42 | 38.9 | 14 | 1 | ANGT_HORSE |

| | | | | | |
|----|----|------|-----|---|------------|
| 34 | 42 | 38.9 | 215 | 1 | RL10_EUGR |
| 35 | 42 | 38.9 | 218 | 1 | RL10_DROME |
| 36 | 42 | 38.9 | 219 | 1 | RL10_BOMMA |
| 37 | 42 | 38.9 | 258 | 1 | GRAX_RAT |
| 38 | 42 | 38.9 | 266 | 1 | DHPS_STRP3 |
| 39 | 42 | 38.9 | 266 | 1 | DHPS_STRP8 |
| 40 | 42 | 38.9 | 266 | 1 | DHPS_STRP8 |
| 41 | 42 | 38.9 | 288 | 1 | SP4G_BACSU |
| 42 | 42 | 38.9 | 476 | 1 | ANGT_SHEEP |
| 43 | 42 | 38.9 | 505 | 1 | MATK_NUPVA |
| 44 | 42 | 38.9 | 523 | 1 | E2BD_HUMAN |
| 45 | 42 | 38.9 | 523 | 1 | E2BD_RABIT |

ALIGNMENTS

RESULT 1
RO60_HUMAN
ID RO60_HUMAN STANDARD; PRT; 538 AA.
AC P10155; Q92787;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 60-kDa SS-A/Ro ribonucleoprotein (60 kDa Ro protein) (60 kDa
DE ribonucleoprotein Ro) (RoNP) (Ro 60 kDa autoantigen) (Sjogren
DE syndrome type A antigen) (SS-A) (Sjogren syndrome antigen A2).
GN SSA2 OR RO60.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=89071722; PubMed=3200833;
RA Deutscher S.L., Harley J.B., Keene J.D.;
RT "Molecular analysis of the 60-kDa human Ro ribonucleoprotein.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:9479-9483(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89198084; PubMed=2649513;
RX Ben-Chetrit E., Gandy B.J., Tan E.M., Sullivan K.F.;
RT "Isolation and characterization of a cDNA clone encoding the 60-kD
RT component of the human SS-A/Ro ribonucleoprotein autoantigen.";
RL J. Clin. Invest. 83:1284-1292(1989).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RA Buyon J.P., DiDonato F., Tseng C.E., Rashbaum W., Morris A.,
RA Hamel J.C., Chan E.K.L.;
RT "Identification and characterization of an alternative mRNA transcript
RT of the 60-kD SS-A/Ro ribonucleoprotein encoding the N-terminal RNA
RL binding domain alone.";
RN Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RP [4]
RP INTERACTION WITH RIP1.
RP TISSUE=keratinocytes;
RX MEDLINE=20015020; PubMed=10545525;
RX Wang D., Buyon J.P., Zhu W., Chan E.K.L.;
RT "Defining a novel 75-kDa phosphoprotein associated with SS-A/Ro and
RT identification of distinct human autoantibodies.";
RL J. Clin. Invest. 104:1265-1275(1999).
CC -!- FUNCTION: RNA-binding protein that binds to several small
CC cytoplasmic RNA molecules known as Y RNAs. May stabilize these
CC RNAs from degradation.
CC -!- SUBUNIT: Binds RIP1.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=P10155-1; Sequence=Displayed;
CC Name=Short; Synonym=60E2;
CC IsoId=P10155-2; Sequence=VSP_005911, VSP_005912;
CC -!- DISEASE: Sera from patients with systemic lupus erythematosus

```

CC (SLE) often contain antibodies that react with the normal cellular
CC SSA2 protein as if this antigen was foreign.
CC -!- SIMILARITY: Belongs to the Ro 60 kDa family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J04137; AAA35493.1; -.
CC DR EMBL; M25077; AAA35532.1; -.
CC DR EMBL; U44388; AAB81552.1; ALT_TERM.
CC DR EMBL; U44388; AAB81553.1; -.
CC DR PIR; A31760; A31760
CC DR Genew; HGNC:11313; SSA2.
CC DR MIM; 600063; -.
CC DR MIM; 234700; -.
CC DR GO; GO:0030529; C:ribonucleoprotein complex; TAS.
CC DR GO; GO:0003723; P:RNA binding; TAS.
CC DR GO; GO:0006383; P:transcription from Pol III promoter; TAS.
CC DR InterPro; IPR000504; RNA_rec_mot.
CC DR Pfam; PF05731; TROVE; 1.
CC DR Ribonucleoprotein; RNA-binding; Systemic lupus erythematosus; Antigen;
CC KW Alternative splicing.
CC FT VARSPLIC 195 205 LAIVTKYITKG -> KHKIFGKGG (in isoform
CC FT Short)
CC FT VARSPLIC 206 538 Missing (in isoform Short).
CC FT FTid=VSP 005911.
CC FT FTid=VSP 005912.
CC FT CONFLICT 239 239 K -> R (IN REF. 2).
CC FT CONFLICT 515 538 GMLDMCGPDTGALDVIRNFDLMI -> ALQNTLLNKSF
CC FT (IN REF. 2).
CC SQ SEQUENCE 538 AA; 60642 MW; 89C58AF3248DC48C CRC64;

Query Match 100.0%; Score 108; DB 1; Length 538;
Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KARIHPHILIALETYKTHG 20
Db 316 KARIHPHILIALETYKTHG 335

RESULT 2
RO60_XENLA STANDARD; PRT; 538 AA.
AC P42700;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 60-kDa SS-A/Ro ribonucleoprotein (60 kDa Ro protein) (60 kDa
DE ribonucleoprotein Ro) (RoRNP).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo, and Ovary;
RX MEDLINE=93348251; PubMed=7688474;
RA O'Brien C.A., Margelot K., Wolin S.L.;
RT "Xenopus Ro ribonucleoproteins: members of an evolutionarily
RT conserved class of cytoplasmic ribonucleoproteins."
RL Proc. Natl. Acad. Sci. U.S.A. 90:7250-7254(1993).
CC -!- FUNCTION: RNA-binding protein that binds to several small
CC cytoplasmic RNA molecules known as Y RNAs. May stabilize these
CC RNAs from degradation.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the Ro 60 kDa family.

```

```

CC -!- SIMILARITY: Belongs to the Ro 60 kDa family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L15430; AAC38001.1; -.
CC DR PIR; I51560; I51560.
CC DR InterPro; IPR000504; RNA_rec_mot.
CC DR InterPro; IPR008858; TROVE.
CC DR Pfam; PF05731; TROVE; 1.
CC DR Ribonucleoprotein; RNA-binding.
CC SQ SEQUENCE 538 AA; 60688 MW; 548C17B0AE9BFD6 CRC64;

Query Match 93.5%; Score 101; DB 1; Length 538;
Best Local Similarity 90.0%; Pred. No. 1.9e-09;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KARIHPHILIALETYKTHG 20
Db 316 KARIHPHILIALETYKTHG 335

RESULT 3
RO60_MOUSE STANDARD; PRT; 538 AA.
AC O08848; Q9QYD8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 60-kDa SS-A/Ro ribonucleoprotein (60 kDa Ro protein) (60 kDa
DE ribonucleoprotein Ro) (RoRNP).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=97266462; PubMed=9112230;
RA Wang D., Buyon J.P., Chan E.K.L.;
RT "Cloning and expression of mouse 60 kDa ribonucleoprotein SS-A/Ro."
RL Mol. Biol. Rep. 23:205-210(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RX MEDLINE=21040327; PubMed=11196703;
RA Kaufman K.M., Farris A.D., Gross J.K., Kirby M.Y., Harley J.B.;
RT "Characterization and genomic sequence of the murine 60 kD Ro gene."
RL Genes Immun. 1:265-270(2000).
RN [3]
RP SEQUENCE OF 82-538 FROM N.A.
RX MEDLINE=99255038; PubMed=1023459;
RA Scofield R.H., Kaufman K.M., Baber U., James J.A., Harley J.B.,
RA Kurien B.T.;
RT "Immunization of mice with human 60-kD Ro peptides results in epitope
RT spreading if the peptides are highly homologous between human and
RT mouse."
RL Arthritis Rheum. 42:1017-1024(1999).
CC -!- FUNCTION: RNA-binding protein that binds to several small
CC cytoplasmic RNA molecules known as Y RNAs. May stabilize these
CC RNAs from degradation.
CC -!- SUBUNIT: Binds RPL11 (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: Highest in brain, followed by lung, muscle,
CC kidney and heart. Lower levels are found in testis, liver and
CC spleen.
CC -!- SIMILARITY: Belongs to the Ro 60 kDa family.

```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U66843; AAC53142.1; --
CC DR EMBL; AF065398; AAF19049.1; --
CC DR EMBL; AF042139; AAC15667.1; --
CC DR MGD; MGI:106652; S8a2.
CC DR InterPro; IPR008858; TROVE.
CC DR Pfam; PF05731; TROVE; 1.
CC KW Ribonucleoprotein; RNA-binding.
CC FT CONFLICT 32 33 RL -> V (IN REF. 2).
CC FT CONFLICT 458 458 D -> G (IN REF. 2).
CC FT CONFLICT 465 465 V -> I (IN REF. 2).
CC FT CONFLICT 465 465 V -> I (IN REF. 2).
CC SQ SEQUENCE 538 AA; 60123 MW; 7850DA35D1726BDA CRC64;

Query Match 91.7%; Score 99; DB 1; Length 538;
Best Local Similarity 85.0%; Pred. No. 4.1e-09;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPHILIAETKTH 20
DQ 316 KARIHPHVLIAETVRAGH 335

RESULT 4
RL10 TOBAC STANDARD; PRT; 150 AA.
AC Q40592;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 60S ribosomal protein L10 (QM protein homolog) (Fragment).
GN RPL10.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. 19;
RA Meyer Y.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the L10e family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z14083; CAA78461.1; --
CC DR PIR; S44144; S44144.
CC DR InterPro; IPR001197; Ribosomal_L10E.
CC DR Pfam; PF00826; Ribosomal_L10e; 1.
CC DR PROSITE; PS01257; RIBOSOMAL_L10E; 1.
CC KW Ribosomal protein.
CC FT NON TER 1
CC SQ SEQUENCE 150 AA; 16517 MW; F265B9D9DFD9B4E CRC64;

Query Match 42.6%; Score 46; DB 1; Length 150;
Best Local Similarity 60.0%; Pred. No. 1.2;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPHIL 10
```

```
Db 17 RVRVHPFHL 26

RESULT 5
RL02 ORYSA STANDARD; PRT; 218 AA.
AC P45636;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 60S ribosomal protein L10-2 (Putative tumor suppressor SG12).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Indica-IR36;
RA Kim J.K.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the L10e family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X81692; CAA57340.1; --
CC DR PIR; S49596; S49596.
CC DR Gramene; P45636; --
CC DR InterPro; IPR001197; Ribosomal_L10E.
CC DR Pfam; PF00826; Ribosomal_L10e; 1.
CC DR TIGRFAMs; TIGR00279; L10e; 1.
CC DR PROSITE; PS01257; RIBOSOMAL_L10E; 1.
CC KW Ribosomal protein.
CC SQ SEQUENCE 218 AA; 24486 MW; 0A053A22AA029107 CRC64;

Query Match 42.6%; Score 46; DB 1; Length 218;
Best Local Similarity 60.0%; Pred. No. 1.7;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPHIL 10
DQ 87 RVRVHPFHL 96

RESULT 6
RL03 ORYSA STANDARD; PRT; 219 AA.
AC Q40649; Q40717;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 60S ribosomal protein L10-3 (QM/R22).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. INDICA GUANG-LU-AI NO.4;
RA Zong H.;
RL Thesis (1996), Fudan University, China.
RN [2]
RP SEQUENCE OF 77-219 FROM N.A.
RC STRAIN=cv. Indica-IR36; TISSUE=Seed;
RA Kim J.K.;
```

```

Best Local Similarity 60.0%; Pred. No. 1.7;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPEHIL 10
Db 88 RVRVHPFHL 97
:|:|:|:|:|
88 RVRVHPFHL 97

RESULT 8
RL10_EUPES
IID RL10_EUPES STANDARD; PRT; 220 AA.
AC Q9MSW7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 60S ribosomal protein L10.
GN RPL10.
OS Euphorbia esula (leafy spurge).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eucots; rosids;
OC eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae; Euphorbieae;
OC Euphorbia
OC NCBI_TaxID=3993;
OX
RN
SEQUENCE FROM N.A.
Anderson J.V., Horvath D.P.;
"Identification of mRNAs expressed in underground adventitious buds of
Euphorbia esula (leafy spurge).";
Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
-!- SIMILARITY: Belongs to the L10e family of ribosomal proteins.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/
or send an email to license@isb-sib.ch).
-----
EMBL; AF227620; AAF34765.1; -
InterPro; IPR001197; Ribosomal_L10E.
Pfam; PF00826; Ribosomal_L10e; 1.
TIGRFAMs; TIGR00279; L10e; 1.
PROSITE; PS01257; RIBOSOMAL_L10E; 1.
Ribosomal protein.
SEQUENCE 220 AA; 24947 MW; E899467B83A56D85 CRC64;

Query Match 42.6%; Score 46; DB 1; Length 220;
Best Local Similarity 60.0%; Pred. No. 1.7;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPEHIL 10
Db 88 RVRVHPFHL 97
:|:|:|:|:|
88 RVRVHPFHL 97

RESULT 9
RL10_MAIZE
IID RL10_MAIZE STANDARD; PRT; 220 AA.
AC P45633;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 60S ribosomal protein L10 (QM protein homolog).
GN RPL10.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; panicoideae; Andropogoneae; Zea.
OC NCBI_TaxID=4577;
OX
RN
SEQUENCE FROM N.A.
STRAIN=cv. B73.

```



```
RX MEDLINE=94362669; PubMed=8081358;
RA Neill J.D.;
RT "Extreme evolutionary conservation of QM, a novel c-Jun associated
RL transcription factor.";
RL Hum. Mol. Genet. 3:723-728(1994).
CC -1- SIMILARITY: Belongs to the L10e family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U06108; AAA17419.1; --
CC PIR; T02068; T02068.
CC MaizeDB; 77933; --
CC InterPro; IPR001197; Ribosomal_L10E.
CC Pfam; PF00826; Ribosomal_L10e; 1.
CC TIGRFAMs; TIGR00279; L10e; 1.
CC PROSITE; PS01257; RIBOSOMAL_L10E; 1.
CC Ribosomal protein.
CC KW RIBOSOMAL protein.
CC SEQUENCE 220 AA; 24919 MW; 5B338061E25893E8 CRC64;
Query Match 42.6%; Score 46; DB 1; Length 220;
Best Local Similarity 60.0%; Pred. No. 1.7;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 KARIHPFHIL 10
DB 88 RVRVHPFVL 97
:|:|:|:|:|
SEQUENCE FROM N.A.
MEDLINE=95247791; PubMed=7730379;
RX Tron T., Yang M., Dick F.A., Schmitt M.E., Trumpower B.L.;
RA "OSR1, an essential yeast gene with a genetic relationship to a
RT subunit of the mitochondrial cytochrome bcl complex, is homologous to
RT a gene implicated in eukaryotic cell differentiation.";
RL J. Biol. Chem. 270:9961-9970(1995).
RN [2]
RN SEQUENCE FROM N.A.
MEDLINE=96381247; PubMed=8789260;
RX Koller H.T., Klade T., Ellinger A., Breitenbach M.;
RA "The yeast growth control gene GRCS is highly homologous to the
RT mammalian putative tumor suppressor gene QM.";
RL Yeast 12:53-65(1996).
RN [3]
RN SEQUENCE FROM N.A.
STRAIN=S288C / AB972;
RX MEDLINE=97333267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansorge W.,
RA Benes V., Brueckner M., Delius H., Dubois E., Dueterhoeft A.,
RA Etian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portecelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Schafie M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestarazu L.A., Vandenbol M., Verbaaselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:87-90(1997).
RN [4]
RN IDENTIFICATION AS A RIBOSOMAL PROTEIN.
RX MEDLINE=97446434; PubMed=9301022;
RA Nika J., Erickson F.L., Hannig E.M.;
RT "Ribosomal protein L9 is the product of GRCS, a homolog of the
RT putative tumor suppressor QM in S. cerevisiae.";
RL Yeast 13:1155-1166(1997).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the L10e family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF180758; AAD56018.1; --
CC InterPro; IPR001197; Ribosomal_L10E.
CC Pfam; PF00826; Ribosomal_L10e; 1.
CC TIGRFAMs; TIGR00279; L10e; 1.
CC PROSITE; PS01257; RIBOSOMAL_L10E; 1.
CC Ribosomal protein.
CC KW RIBOSOMAL protein.
CC SEQUENCE 220 AA; 25047 MW; 4BF7A938C5BB51B6 CRC64;
Query Match 42.6%; Score 46; DB 1; Length 220;
```

```

DR EMBL; X78887; CAA55485.1; -.
DR EMBL; Z73247; CAA97632.1; -.
DR PIR; A57296; A57296.
DR PDB; 1K5Y; 22-MAY-02.
DR GenOnline; 142137; -.
DR SGD; S0004065; RPL10.
DR GO; GO:0000027; Ribosomal large subunit assembly and mainte. .; IMP.
DR InterPro; IPR001197; Ribosomal L10E.
DR Pfam; PF00826; Ribosomal L10e; 1.
DR TIGRFAMs; TIGR00279; L10e; 1.
DR PROSITE; PS01257; RIBOSOMAL_L10E; 1.
DR Ribosomal protein; Repeat; 3D-structure.
FT DOMAIN 167 216 2 X 15-17 AA APPROXIMATE REPEATS.
FT REPEAT 167 181 1.
FT REPEAT 200 216 2.
FT MUTAGEN 194 194 G->D; IN QGR1-1; SYNTHETIC LETHAL.
SQ SEQUENCE 221 AA; 25361 MW; 6FF976A3800F347D CRC64;

Query Match 42.6%; Score 46; DB 1; Length 221;
Best Local Similarity 60.0%; Pred. No. 1.8;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPHIL 10
Db 88 RVRVHPHVL 97

RESULT 12
RL10 PINTA
ID RL10 PINTA STANDARD; PRT; 228 AA.
AC 022431;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 60S ribosomal protein L10 (Wilm's tumor suppressor homolog).
GN RPL10 OR LP20.
OS Pinus taeda (loblolly pine).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=3352;
RN [1]
RP SEQUENCE FROM N.A.
RA Padmanabhan V., Dias M.A.D.L., Newton R.J.;
RT "Isolation and characterization of a water deficit stress repressible
RT Wilm's tumor suppressor homolog from loblolly pine (Pinus taeda).";
RT Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the L10e family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF013804; AA866347.1; -.
DR PIR; T07957; T07957.
DR InterPro; IPR001197; Ribosomal L10E.
DR Pfam; PF00826; Ribosomal L10e; 1.
DR TIGRFAMs; TIGR00279; L10e; 1.
DR PROSITE; PS01257; RIBOSOMAL_L10E; 1.
DR Ribosomal protein.
SQ SEQUENCE 228 AA; 25976 MW; E1E5D5B905A07BCA CRC64;

Query Match 42.6%; Score 46; DB 1; Length 228;
Best Local Similarity 60.0%; Pred. No. 1.8;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPHIL 10
Db 88 RVRVHPHVL 97

```

```

RESULT 13
DP01_LACLA
ID DP01_LACLA STANDARD; PRT; 877 AA.
AC Q9CDS1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase I (EC 2.7.7.7) (POL 1).
GN POLA OR LL2142.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=1137471;
RA Boletín A., Wincker P., Mauger S., Jailion O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403".
RL Genome Res. 11:731-753(2001).
CC -!- FUNCTION: In addition to polymerase activity, This DNA polymerase
CC exhibits 3' to 5' and 5' to 3' exonuclease activity (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC -!- SUBUNIT: Single-chain monomer with multiple functions.
CC -!- SIMILARITY: Belongs to the DNA polymerase type-A family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE006442; AAK06240.1; -.
DR PIR; F86892; F86892.
DR HSSP; P52026; 1XWL.
DR InterPro; IPR002562; 3_5_exonuclease.
DR InterPro; IPR008918; 5_3_exo_C.
DR InterPro; IPR002421; 5_3_exonuclease.
DR InterPro; IPR001098; DNA_pol.
DR InterPro; IPR002298; DNA_pol.
DR InterPro; IPR000513; Exo_N_I.
DR Pfam; PF02739; 5_3_exonuc_N; 1.
DR Pfam; PF01367; 5_3_exonuclease; 1.
DR Pfam; PF00476; DNA_pol_A; 1.
DR PRINTS; PR00868; DNAPOLI.
DR SMART; SM00474; 35EXOC; 1.
DR SMART; SM00475; 53EXOC; 1.
DR SMART; SM00279; HHZ; 1.
DR SMART; SM00482; POLAC; 1.
DR TIGRFAMs; TIGR00593; pola; 1.
DR PROSITE; PS00447; DNA_POLYMERASE_A; 1.
DR Transferase; DNA-directed DNA polymerase; DNA replication; DNA repair;
DR Hydrolase; Exonuclease; DNA-binding; Complete proteome.
SQ SEQUENCE 877 AA; 98732 MW; A5C2BEB92FF98FB2 CRC64;

Query Match 42.6%; Score 46; DB 1; Length 877;
Best Local Similarity 56.2%; Pred. No. 7.9;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 RIHPHILIAETKYT 18
Db 54 RIEPTHVLIADFAGKT 69

RESULT 14
DP01_LACLC
ID DP01_LACLC STANDARD; PRT; 877 AA.

```


Thu May 20 11:48:33 2004

us-10-005-684-6.rsp

Page 8

Db 88 RIRVPHFVL 97

Search completed: May 19, 2004, 15:44:30
Job time : 10.5833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 15:39:22 ; Search time 14.5833 Seconds
(without alignments)
131.920 Million cell updates/sec

Title: US-10-005-684-6
Perfect score: 108
Sequence: 1 KARIHPHILIALETYKTH 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--|
| 1 | 108 | 100.0 | 538 | 2 A31760 | Ro/SS-A complex, 60K ribonucleoprotein - human |
| 2 | 101 | 93.5 | 538 | 2 I51560 | ribonucleoprotein |
| 3 | 48 | 44.4 | 997 | 2 I5243 | hypothetical prote |
| 4 | 47 | 43.5 | 188 | 2 C90112 | 60S ribosomal prot |
| 5 | 46 | 42.6 | 143 | 2 S19224 | ribosomal protein |
| 6 | 46 | 42.6 | 150 | 2 S44144 | ribosomal protein |
| 7 | 46 | 42.6 | 218 | 2 S49596 | ribosomal protein |
| 8 | 46 | 42.6 | 220 | 2 T02068 | probable transcrip |
| 9 | 46 | 42.6 | 221 | 2 A57296 | ribosomal protein |
| 10 | 46 | 42.6 | 228 | 2 T07957 | probable ribosomal |
| 11 | 46 | 42.6 | 877 | 2 F86892 | DNA-directed DNA p |
| 12 | 45 | 41.7 | 184 | 2 F96691 | probable 60S ribos |
| 13 | 45 | 41.7 | 220 | 2 E86277 | hypothetical prote |
| 14 | 45 | 41.7 | 221 | 2 A86396 | 60S ribosomal prot |
| 15 | 45 | 41.7 | 252 | 1 E69839 | conserved hypotet |
| 16 | 45 | 41.7 | 338 | 2 T41021 | ser-thr protein K1 |
| 17 | 45 | 41.7 | 643 | 2 T19225 | Ro autoantigen 60K |
| 18 | 45 | 41.7 | 877 | 2 A32949 | DNA-directed DNA p |
| 19 | 45 | 41.7 | 877 | 2 E95003 | DNA polymerase I |
| 20 | 45 | 41.7 | 877 | 2 H97875 | DNA-directed DNA p |
| 21 | 44 | 40.7 | 184 | 2 S74232 | bidirectional hydr |
| 22 | 44 | 40.7 | 304 | 2 G83820 | GTP-binding protei |
| 23 | 44 | 40.7 | 633 | 2 T04179 | hypothetical prote |
| 24 | 44 | 40.7 | 1099 | 2 H83210 | probable phospholi |
| 25 | 43 | 39.8 | 210 | 2 A48226 | ribosomal protein |
| 26 | 43 | 39.8 | 214 | 1 JC4911 | ribosomal protein |
| 27 | 43 | 39.8 | 214 | 2 A42735 | ribosomal protein |
| 28 | 43 | 39.8 | 214 | 2 JC2013 | ribosomal protein |
| 29 | 43 | 39.8 | 272 | 2 AF0595 | conserved hypotet |

| | | | | | | |
|----|------|------|------|---|--------|--------------------|
| 30 | 43 | 39.8 | 292 | 2 | T06684 | hypothetical prote |
| 31 | 42.5 | 39.4 | 791 | 2 | E35216 | FPD5 protein - fow |
| 32 | 42.5 | 39.4 | 887 | 2 | B56679 | probable reverse t |
| 33 | 42 | 38.9 | 14 | 2 | A01250 | angiotensin I prec |
| 34 | 42 | 38.9 | 15 | 2 | A60834 | angiotensin I prec |
| 35 | 42 | 38.9 | 222 | 2 | S74596 | hypothetical prote |
| 36 | 42 | 38.9 | 258 | 2 | I56220 | tryptase 2 - rat |
| 37 | 42 | 38.9 | 288 | 2 | S18438 | sporulation protei |
| 38 | 42 | 38.9 | 394 | 2 | T08757 | probable translati |
| 39 | 42 | 38.9 | 476 | 1 | JC2318 | angiotensin precu |
| 40 | 42 | 38.9 | 523 | 2 | S42727 | translation initia |
| 41 | 42 | 38.9 | 524 | 2 | S71961 | translation initia |
| 42 | 42 | 38.9 | 544 | 2 | A55146 | guanine nucleotide |
| 43 | 42 | 38.9 | 694 | 2 | D87330 | senary box/GGDEF |
| 44 | 42 | 38.9 | 1084 | 2 | T13173 | RNA-directed DNA p |
| 45 | 42 | 38.9 | 1304 | 1 | A46546 | leukocyte common a |

ALIGNMENTS

RESULT 1

A31760
Ro/SS-A complex, 60K ribonucleoprotein - human
C/Species: Homo sapiens (man)
C/Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 22-Oct-1999
C/Accession: A31760; A30596
R/Deutscher, S.L.; Harley, J.B.; Keene, J.D.
Proc. Natl. Acad. Sci. U.S.A. 85, 9479-9483, 1988
A/Title: Molecular analysis of the 60-kDa human Ro ribonucleoprotein.
A/Reference number: A31760; MUID:89071722; PMID:3200833
A/Accession: A31760
A/Molecule type: mRNA
A/Residues: 1-538 <DEU>
A/Cross-references: GB:J04137; NID:gl77782; PIDN:AAA35493.1; PID:gl77783
R/Ben-Chetrit, E.; Gandy, B.J.; Tan, E.M.; Sullivan, K.F.
J. Clin. Invest. 83, 1284-1292, 1989
A/Title: Isolation and characterization of a cDNA clone encoding the 60-kD component of
A/Reference number: A30596; MUID:89198084; PMID:2649513
A/Accession: A30596
A/Molecule type: mRNA
A/Residues: 1-238, 'R', 240-292, 'DV', 295-300, 'A', 302-514, 'ALQNTLLNKSF' <BEN>
A/Cross-references: GB:M25077; NID:G387656
A/Note: the sequence is revised in GenBank entry HUMANTARNP, release 111.0, (PID:G387657)

C/Genetics:
A/Gene: GDB:SSA2
A/Cross-references: GDB:355563; OMIM:600063
A/Map position: 1q31-1q31
C/Keywords: DNA binding; zinc finger
F:305-323/Region: zinc finger CCHH motif

Query Match 100.0%; Score 108; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KARIHPHILIALETYKTH 20
|||
Db 316 KARIHPHILIALETYKTH 335

RESULT 2

I51560
ribonucleoprotein - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C/Accession: I51560
R/O'Brien, C.A.; Margelot, K.; Wolin, S.L.
Proc. Natl. Acad. Sci. U.S.A. 90, 7250-7254, 1993
A/Title: Xenopus Ro ribonucleoproteins: members of an evolutionarily conserved class of
A/Reference number: A48294; MUID:93348251; PMID:7688474
A/Accession: I51560
A/Status: preliminary; translated from GB/EMBL/DDBT
A/Molecule type: mRNA

```

A;Residues: 1-538 <OXB>
A;Cross-references: GB:L15430; NID:g295535; PIDN:AAC38001.1; PID:g295536

Query Match      93.5%; Score 101; DB 2; Length 538;
Best Local Similarity 90.0%; Pred. No. 7.8e-09;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPFHILIALETYKTGH 20
    |||||:|||||:|||||
Db 316 KARIHPFHILVALETYKXGH 335

RESULT 3
TI5243
hypothetical protein T05E7.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999
C;Accession: TI5243
R;Rohlfing, T.; Wohlmann, P.; Biewald, T.
submitted to the EMBL Data Library, May 1997
A;Description: The sequence of C. elegans cosmid T05E7.
A;Reference number: Z18315
A;Accession: TI5243
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-997 <ROH>
A;Cross-references: EMBL:AF003150; NID:g2088791; PID:g2088794; PIDN:AAB54215.1; GSPDB:GN
A;Experimental source: strain Bristol N2; clone T05E7
C;Genetics:
A;Gene: CESP:T05E7.3
A;Map position: 1
A;Introns: 124/3; 183/2; 245/3; 298/3; 445/3; 597/3; 640/3; 669/1; 740/2; 795/1; 847/3;
C;Superfamily: Caenorhabditis elegans hypothetical protein T05E7.3

Query Match      44.4%; Score 48; DB 2; Length 997;
Best Local Similarity 56.2%; Pred. No. 11;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 RIHPFHILIALETYKT 18
    :|||||:|:|:|
Db 771 KIVPFHILTSTHIT 786

RESULT 4
C90112
60S ribosomal protein L10 [imported] - Giardia theta nucleomorph
C;Species: nucleomorph Giardia theta
A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 15-Jun-2001
C;Accession: C90112
R;Doughas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Re
Nature 410, 1091-1096, 2001
A;Title: The highly reduced genome of an enslaved algal nucleus.
A;Reference number: A99082; MUID:11323671; PMID:11323671
A;Accession: C90112
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-188 <DOU>
A;Cross-references: GB:AJ010592; NID:g12580744; PIDN:CAC27062.1; GSPDB:GN00151
C;Genetics:
A;Map position: 2
A;Genome: nucleomorph
C;Superfamily: rat ribosomal protein L10
C;Keywords: nucleomorph

Query Match      43.5%; Score 47; DB 2; Length 188;
Best Local Similarity 70.0%; Pred. No. 2.5;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPFHIL 10
    |||||:|:|
Db 88 KIRIHPFHI 97

RESULT 5
S19224
ribosomal protein L10.e, cytosolic - rice (fragment)
C;Species: Oryza sativa (rice)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 18-Aug-2000
C;Accession: S19224
R;Kim, J.K.
submitted to the EMBL Data Library, February 1992
A;Description: A rice novel gene.
A;Reference number: S19224
A;Accession: S19224
A;Molecule type: mRNA
A;Residues: 1-143 <KIM>
A;Cross-references: EMBL:X64621; NID:g203111; PIDN:CAA45905.1; PID:g20312
C;Superfamily: rat ribosomal protein L10
C;Keywords: cytosol; protein biosynthesis; ribosome

Query Match      42.6%; Score 46; DB 2; Length 143;
Best Local Similarity 60.0%; Pred. No. 2.8;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPFHIL 10
    :|:|:|:|
Db 12 RVRVHPFHL 21

RESULT 6
S44144
ribosomal protein L10.e, cytosolic - common tobacco (fragment)
N;Alternate names: WILM's tumor-related protein HUMQM
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 06-Feb-1995 #sequence_revision 06-Dec-1996 #text_change 18-Aug-2000
C;Accession: S44144
R;Meyer, Y.
submitted to the EMBL Data Library, July 1992
A;Description: Growth-related gene expression in Nicotiana tabacum mesophyll protoplasts.
A;Reference number: S29905
A;Accession: S44144
A;Molecule type: mRNA
A;Residues: 1-150 <MEY>
A;Cross-references: EMBL:Z14083; NID:g473103; PIDN:CAA78461.1; PID:g473104
C;Superfamily: rat ribosomal protein L10
C;Keywords: cytosol; protein biosynthesis; ribosome

Query Match      42.6%; Score 46; DB 2; Length 150;
Best Local Similarity 60.0%; Pred. No. 2.9;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPFHIL 10
    :|:|:|:|
Db 17 RVRVHPFHL 26

RESULT 7
S49596
ribosomal protein L10.e, cytosolic - rice
C;Species: Oryza sativa (rice)
C;Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 18-Aug-2000
C;Accession: S49596
R;Kim, J.K.
submitted to the EMBL Data Library, November 1994
A;Description: Isolation and characterization of two rice genes encoding a putative tumor
A;Reference number: S49596
A;Accession: S49596
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-218 <KIM>
A;Cross-references: EMBL:X81692; NID:g805003; PIDN:CAA57340.1; PID:g575357
C;Genetics:
A;Introns: 4/1; 170/2; 197/3
C;Superfamily: rat ribosomal
C;Keywords: cytosol; protein biosynthesis; ribosome

```

R;Koller, H.T.; Klade, T.; Ellinger, A.; Breitenbach, M.
Yeast 12, 53-65, 1996
A;Title: The yeast growth control gene GRC5 is highly homologous to the mammalian putative growth control gene GRC5
A;Reference number: S61714; MUID:96381247; PMID:8789260
A;Accession: S61714
A;Molecule type: DNA
A;Residues: 1-221 <KOF>
A;Cross-references: EMBL:X78887; NID:gl679862; PIDN:CAA55485.1; PID:g747904
R;Pohl, T.M.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64899
A;Accession: S64907
A;Molecule type: DNA
A;Residues: 1-221 <POH>
A;Cross-references: EMBL:Z73247; NID:gl360433; PIDN:CAA97632.1; PID:gl360434; MIPS:YLR075w
A;Map position: 12R
C;Superfamily: rat ribosomal protein L10
C;Keywords: cytosol; protein biosynthesis; ribosome

Query Match 42.6%; Score 46; DB 2; Length 221;
Best Local Similarity 60.0%; Pred. No. 4.4;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPFHIL 10
:|||||:
DB 88 RRVVHPFVL 97

RESULT 10
T07957
probable ribosomal protein L10.e - loblolly pine
C;Species: Pinus taeda (loblolly pine)
C;Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 18-Aug-2000
R;Padmanabhan, V.; Dias, M.A.; Newton, R.J.
submitted to the EMBL Data Library, July 1997
A;Description: Isolation and characterization of a water deficit stress repressible WilD
A;Reference number: Z18245
A;Accession: T07957
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-228 <PAD>
A;Cross-references: EMBL:AF013804; NID:g2317761; PIDN:AAE66347.1; PID:g2317762
C;Genetics:
A;Gene: lp20
C;Superfamily: rat ribosomal protein L10

Query Match 42.6%; Score 46; DB 2; Length 228;
Best Local Similarity 60.0%; Pred. No. 4.6;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPFHIL 10
:|||||:
DB 88 RRVVHPFVL 97

RESULT 11
F86892
DNA-directed DNA polymerase (EC 2.7.7.7) [imported] - Lactococcus lactis subsp. lactis
N;Alternate names: DNA polymerase I
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
R;Bolotin, A.; Wincker, P.; Manger, S.; Jaillon, O.; Malarre, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp
A;Reference number: AB6625; MUID:21235186; PMID:11337471
A;Accession: F86892
A;Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-877 <STO>
 A:Cross-references: GB:AE005176; PID:g12725200; PIDN:AAK06240.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: polA
 C:Superfamily: DNA-directed DNA polymerase I
 C:Keywords: nucleotidyltransferase

Query Match 42.6%; Score 46; DB 2; Length 877;
 Best Local Similarity 56.2%; Pred. No. 20;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 RIHPHILIALEYTKT 18
 |||:||||:
 DB 54 RIETPHVLIAFDAGKT 69

RESULT 12
 E96691
 Probable 60S ribosomal protein L10 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 24-Aug-2001
 C:Accession: F96691
 R:Theologias: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: F96691
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-184 <STO>
 A:Cross-references: GB:AE005173; NID:g11054579; PIDN:AAG27854.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: T1217.3
 C:Superfamily: rat ribosomal protein L10

Query Match 41.7%; Score 45; DB 2; Length 184;
 Best Local Similarity 60.0%; Pred. No. 5.3;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPFHIL 10
 :|||:
 DB 51 RIRVHPFVL 60

RESULT 13
 E86277
 hypothetical protein F14117.9 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 C:Accession: E86277
 R:Theologias: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: E86277
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-220 <STO>
 A:Cross-references: GB:AE005172; NID:g7262674; PIDN:AAF43932.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: rat ribosomal protein L10

Query Match 41.7%; Score 45; DB 2; Length 220;
 Best Local Similarity 60.0%; Pred. No. 6.5;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPFHIL 10
 :|||:
 DB 88 RIRVHPFVL 97

RESULT 14
 A86396
 60S ribosomal protein L10 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: A86396
 R:Theologias: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: A86396
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-221 <STO>
 A:Cross-references: GB:AE005172; NID:g4262180; PIDN:AAD14497.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1
 C:Superfamily: rat ribosomal protein L10

Query Match 41.7%; Score 45; DB 2; Length 221;
 Best Local Similarity 60.0%; Pred. No. 6.5;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPFHIL 10
 :|||:
 DB 88 RIRVHPFVL 97

RESULT 15
 E69839
 conserved hypothetical protein yitD - Bacillus subtilis
 C:Species: Bacillus subtilis
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 31-Mar-2003
 C:Accession: E69839
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi
 A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallero
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.P.;
 Koetter, P.; Konningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A.; Laub, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauesl,
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
 A.; Authors: Schleicher, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,
 T.; Winters, P.; Wipat, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 A.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Yata, K.; Yoshida, K.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: E69839
 B:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-252 <KUN>
 A:Cross-references: GB:Z99109; GB:AL009126; NID:G2633260; PIDN:CAB12935.1; PID:G2633431
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: YtCD
 C:Superfamily: (2R)-phospho-3-sulfolactate synthase

| | | | | |
|-----------------------|-------|-----------------|---------------|-------------------|
| Query Match | 41.7% | Score 45; | DB 1; | Length 252; |
| Best Local Similarity | 64.3% | Pred. No. 7.5; | | |
| Matches | 9; | Conservative 1; | Mismatches 4; | Indels 0; Gaps 0; |

QY 6 PFHILIALETYKTG 19
 ||| ||||| :
 Db 231 PFHDAIALETIRLG 244

Search completed: May 19, 2004, 15:47:11
 Job time : 15.5833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 15:27:51 ; Search time 45 Seconds
(without alignments)
123.672 Million cell updates/sec

Title: US-10-005-684-6

Perfect score: 108

Sequence: 1 KARIHPFHILIALEYTKTGH 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1145568 seqs, 278261457 residues

Total number of hits satisfying chosen parameters: 1145568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCF_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCFUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------------------|--------------------|
| 1 | 108 | 100.0 | 20 | 14 US-10-005-684-6 | Sequence 6, Appli |
| 2 | 108 | 100.0 | 552 | 14 US-10-288-858-1 | Sequence 1, Appli |
| 3 | 108 | 100.0 | 553 | 14 US-10-288-858-2 | Sequence 2, Appli |
| 4 | 108 | 100.0 | 553 | 14 US-10-288-858-6 | Sequence 6, Appli |
| 5 | 51 | 47.2 | 17 | 15 US-10-376-121A-58 | Sequence 58, Appli |
| 6 | 46 | 42.6 | 203 | 12 US-10-425-114-47741 | Sequence 47741, A |
| 7 | 46 | 42.6 | 205 | 12 US-10-425-114-47741 | Sequence 47741, A |
| 8 | 46 | 42.6 | 222 | 12 US-10-424-599-273820 | Sequence 273820, A |
| 9 | 46 | 42.6 | 223 | 12 US-10-424-599-273820 | Sequence 273820, A |
| 10 | 46 | 42.6 | 223 | 12 US-10-424-599-273817 | Sequence 273817, A |
| 11 | 46 | 42.6 | 223 | 12 US-10-424-599-273818 | Sequence 273818, A |
| 12 | 46 | 42.6 | 223 | 12 US-10-424-599-273818 | Sequence 273818, A |
| 13 | 46 | 42.6 | 229 | 12 US-10-424-599-152088 | Sequence 152088, A |
| 14 | 46 | 42.6 | 236 | 12 US-10-424-599-224269 | Sequence 224269, A |
| 15 | 45 | 41.7 | 338 | 15 US-10-369-493-22732 | Sequence 22732, A |

| | | | | | | |
|----|------|------|------|----|-----------------------|---------------------|
| 16 | 45 | 41.7 | 707 | 14 | US-10-032-189-125 | Sequence 125, App |
| 17 | 45 | 41.7 | 877 | 12 | US-10-282-122A-73692 | Sequence 73692, A |
| 18 | 45 | 41.7 | 889 | 9 | US-09-864-761-36605 | Sequence 13236, A |
| 19 | 44 | 40.7 | 95 | 12 | US-10-424-599-2254420 | Sequence 2254420, A |
| 20 | 44 | 40.7 | 112 | 12 | US-10-424-599-225844 | Sequence 225844, A |
| 21 | 44 | 40.7 | 117 | 12 | US-10-424-599-221584 | Sequence 221584, A |
| 22 | 44 | 40.7 | 155 | 12 | US-10-424-599-281988 | Sequence 281988, A |
| 23 | 44 | 40.7 | 633 | 12 | US-10-225-066A-388 | Sequence 388, App |
| 24 | 44 | 40.7 | 633 | 15 | US-10-374-780A-2740 | Sequence 2740, App |
| 25 | 44 | 40.7 | 878 | 12 | US-10-282-122A-72012 | Sequence 72012, A |
| 26 | 44 | 40.7 | 880 | 12 | US-10-282-122A-74355 | Sequence 74355, A |
| 27 | 44 | 40.7 | 2566 | 15 | US-10-080-334-162 | Sequence 162, App |
| 28 | 44 | 40.7 | 2566 | 15 | US-10-080-334-163 | Sequence 163, App |
| 29 | 44 | 40.7 | 2568 | 9 | US-09-866-108-3 | Sequence 3, Appli |
| 30 | 43.5 | 40.3 | 75 | 12 | US-10-424-599-177044 | Sequence 177044, A |
| 31 | 43 | 39.8 | 34 | 9 | US-09-864-761-36605 | Sequence 36605, A |
| 32 | 43 | 39.8 | 97 | 12 | US-10-424-599-157738 | Sequence 157738, A |
| 33 | 43 | 39.8 | 118 | 9 | US-09-764-846-160 | Sequence 160, App |
| 34 | 43 | 39.8 | 118 | 14 | US-10-091-483-160 | Sequence 160, App |
| 35 | 43 | 39.8 | 206 | 14 | US-10-146-473-56 | Sequence 56, Appli |
| 36 | 43 | 39.8 | 214 | 15 | US-10-012-697-1566 | Sequence 1566, App |
| 37 | 43 | 39.8 | 223 | 14 | US-10-029-386-32466 | Sequence 32466, A |
| 38 | 43 | 39.8 | 243 | 15 | US-10-264-049-3449 | Sequence 3449, App |
| 39 | 43 | 39.8 | 372 | 10 | US-09-890-688-62 | Sequence 62, Appli |
| 40 | 43 | 39.8 | 372 | 14 | US-10-181-071-6 | Sequence 6, Appli |
| 41 | 43 | 39.8 | 374 | 9 | US-09-764-846-236 | Sequence 236, App |
| 42 | 43 | 39.8 | 374 | 14 | US-10-091-483-236 | Sequence 236, App |
| 43 | 43 | 39.8 | 722 | 13 | US-10-128-870-23 | Sequence 23, Appli |
| 44 | 43 | 39.8 | 722 | 14 | US-10-131-685-23 | Sequence 23, Appli |
| 45 | 43 | 39.8 | 757 | 14 | US-10-096-578-89 | Sequence 89, Appli |

ALIGNMENTS

RESULT 1
US-10-005-684-6
; Sequence 6, Application US/10005684
; Publication No. US20030100035A1
; GENERAL INFORMATION:
; APPLICANT: Vojdani, Ariosto
; TITLE OF INVENTION: SALIVA IMMUNOASSAY FOR DETECTION OF
; TITLE OF INVENTION: ANTIBODIES FOR AUTOIMMUNE DISEASE
; FILE REFERENCE: IMSC12.005A
; CURRENT APPLICATION NUMBER: US/10/005,684
; CURRENT FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Lupus peptide
US-10-005-684-6

Query Match 100.0%; Score 108; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KARIHPFHILIALEYTKTGH 20
DB 1 KARIHPFHILIALEYTKTGH 20

RESULT 2
US-10-288-858-1
; Sequence 1, Application US/10288858
; Publication No. US20030109001A1
; GENERAL INFORMATION:
; APPLICANT: Burckhardt, Jean
; APPLICANT: Haass, Michael
; APPLICANT: Lehmann, Hans-Peter
; TITLE OF INVENTION: Process for the Recombinant Production of Ribonucleoproteins
; FILE REFERENCE: RDID 0052US

```
; CURRENT APPLICATION NUMBER: US/10/288,858
; CURRENT FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: DE 19931380.6
; PRIOR FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: SSA60 M4-C6
US-10-288-858-1
```

```
Query Match      100.0%; Score 108; DB 14; Length 552;
Best Local Similarity 100.0%; Pred. No. 4.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 KARIHPHILIALETYKTH 20
Db      328 KARIHPHILIALETYKTH 347
```

```
RESULT 3
US-10-288-858-2
; Sequence 2, Application US/10288858
; Publication No. US20030109001A1
; GENERAL INFORMATION:
; APPLICANT: Burckhardt, Jean
; APPLICANT: Haass, Michael
; APPLICANT: Lehmann, Hans-Peter
; TITLE OF INVENTION: Process for the Recombinant Production of Ribonucleoproteins
; FILE REFERENCE: RDID 0052US
; CURRENT APPLICATION NUMBER: US/10/288,858
; CURRENT FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: DE 19931380.6
; PRIOR FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: SSA60 M56
US-10-288-858-2
```

```
Query Match      100.0%; Score 108; DB 14; Length 553;
Best Local Similarity 100.0%; Pred. No. 4.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 KARIHPHILIALETYKTH 20
Db      331 KARIHPHILIALETYKTH 350
```

```
RESULT 4
US-10-288-858-6
; Sequence 6, Application US/10288858
; Publication No. US20030109001A1
; GENERAL INFORMATION:
; APPLICANT: Burckhardt, Jean
; APPLICANT: Haass, Michael
; APPLICANT: Lehmann, Hans-Peter
; TITLE OF INVENTION: Process for the Recombinant Production of Ribonucleoproteins
; FILE REFERENCE: RDID 0052US
; CURRENT APPLICATION NUMBER: US/10/288,858
; CURRENT FILING DATE: 2002-11-06
; PRIOR APPLICATION NUMBER: DE 19931380.6
; PRIOR FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 14
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-288-858-6
```

```
Query Match      100.0%; Score 108; DB 14; Length 553;
Best Local Similarity 100.0%; Pred. No. 4.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 KARIHPHILIALETYKTH 20
Db      331 KARIHPHILIALETYKTH 350
```

```
RESULT 5
US-10-376-121A-58
; Sequence 58, Application US/10376121A
; Publication No. US20030216544A1
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS OF
; AUTOANTIBODIES
; NUMBER OF SEQUENCES: 218
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: Suite 2000, 1201 West Peachtree Street, N.E.
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/10/376,121A
; APPLICATION NUMBER: US/10/376,121A
; FILING DATE: 27-Mar-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/867,819
; FILING DATE: April 13, 1992
; APPLICATION NUMBER: 07/648,205
; FILING DATE: January 31, 1991
; APPLICATION NUMBER: 07/472,947
; FILING DATE: January 31, 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRFL14CIP(2)DIV(2)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-817-8473
; TELEFAX: (404)-817-8588
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 6..15
; SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-10-376-121A-58
```

```
Query Match      47.2%; Score 51; DB 15; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 KARIHPHI 9
```

```

Db          ||||| 17
            9 KARIHPFHIL 17

RESULT 6
US-10-425-114-47741
; Sequence 47741, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 47741
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4073-004-B7_FLI.pep
US-10-425-114-47741

Query Match          42.6%; Score 46; DB 12; Length 203;
Best Local Similarity 60.0%; Pred. No. 20;
Matches 6; Conservative 1; Indels 0; Gaps 0;

QY          1 KARIHPFHIL 10
            : |||||:
Db          67 RVRVHPFHL 76

RESULT 7
US-10-425-114-64735
; Sequence 64735, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64735
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4606-007-E2_FLI.pep
US-10-425-114-64735

Query Match          42.6%; Score 46; DB 12; Length 205;
Best Local Similarity 60.0%; Pred. No. 20;
Matches 6; Conservative 1; Indels 0; Gaps 0;

QY          1 KARIHPFHIL 10
            : |||||:
Db          112 RVRVHPFHL 121

RESULT 8
US-10-424-599-273820
; Sequence 273820, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 273820
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89280C.1.pep
US-10-424-599-273820

Query Match          42.6%; Score 46; DB 12; Length 222;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY          1 KARIHPFHIL 10
            : |||||:
Db          88 RVRVHPFHL 97

RESULT 9
US-10-424-599-273814
; Sequence 273814, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 273814
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89275C.1.pep
US-10-424-599-273814

Query Match          42.6%; Score 46; DB 12; Length 223;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY          1 KARIHPFHIL 10
            : |||||:
Db          88 RVRVHPFHL 97

RESULT 10
US-10-424-599-273817
; Sequence 273817, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 273817
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89278C.1.pep
US-10-424-599-273817

Query Match 42.6%; Score 46; DB 12; Length 223;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KARIHPFHIL 10
+ : : : : :
Db 88 RVRVHPFHL 97

RESULT 11

US-10-424-599-273818
; Sequence 273818, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 273818
; LENGTH: 223
; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)-(223)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89279C.1.pep
US-10-424-599-273818

Query Match 42.6%; Score 46; DB 12; Length 223;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KARIHPFHIL 10
+ : : : : :
Db 88 RVRVHPFHL 97

RESULT 12

US-10-424-599-273823
; Sequence 273823, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 273823
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(223)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89283C.1.pep
US-10-424-599-273823

Query Match 42.6%; Score 45; DB 12; Length 223;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KARIHPFHIL 10
+ : : : : :
Db 88 RVRVHPFHL 97

RESULT 13

US-10-424-599-152088
; Sequence 152088, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 152088
; LENGTH: 229
; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_108359C.1.pep
US-10-424-599-152088

Query Match 42.6%; Score 46; DB 12; Length 229;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KARIHPFHIL 10
+ : : : : :
Db 95 RVRVHPFHL 104

RESULT 14

US-10-424-599-224269
; Sequence 224269, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 224269
; LENGTH: 236
; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_44543C.1.pcp
US-10-424-599-224569

Query Match 42.6%; Score 46; DB 12; Length 236;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPFHIL 10
Db 104 RVRVHPFVL 113

RESULT 15

US-10-369-493-22732
; Sequence 22732, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22732
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-22732

Query Match 41.7%; Score 45; DB 15; Length 338;
Best Local Similarity 45.0%; Pred. No. 49;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 KARIHPFHILAEITYKTH 20
Db 295 KLQTHFFCSISLETVSKGN 314

Search completed: May 19, 2004, 15:41:17
Job time : 46 secs

Thu May 20 11:48:31 2004

us-10-005-684-6.ra1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 15:23:26 ; Search time 17.5 Seconds
(without alignments)
59.001 Million cell updates/sec

Title: US-10-005-684-6

Perfect score: 108

Sequence: 1 KARIHPHILIALEYTKGH 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pep:*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB.pep:*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB.pep:*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB.pep:*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pep:*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 51 | 47.2 | 17 | 4 | US-08-475-955-58 |
| 2 | 51 | 47.2 | 21 | 4 | US-08-716-249-11 |
| 3 | 47 | 43.5 | 8 | 3 | US-08-160-604-87 |
| 4 | 47 | 43.5 | 16 | 3 | US-08-160-604-78 |
| 5 | 46 | 42.6 | 220 | 1 | US-08-033-797-2 |
| 6 | 46 | 42.6 | 220 | 1 | US-08-472-265-2 |
| 7 | 46 | 42.6 | 220 | 1 | US-08-472-263-2 |
| 8 | 44 | 40.7 | 2568 | 4 | US-09-866-108A-3 |
| 9 | 43 | 39.8 | 214 | 1 | US-08-033-797-3 |
| 10 | 43 | 39.8 | 214 | 1 | US-08-472-263-3 |
| 11 | 43 | 39.8 | 214 | 1 | US-08-472-263-3 |
| 12 | 43 | 39.8 | 722 | 4 | US-09-105-058C-23 |
| 13 | 43 | 39.8 | 722 | 4 | US-09-177-650-89 |
| 14 | 43 | 39.8 | 757 | 4 | US-09-177-650-89 |
| 15 | 43 | 39.8 | 844 | 4 | US-09-813-148-4 |
| 16 | 43 | 39.8 | 844 | 4 | US-09-590-304-4 |
| 17 | 43 | 39.8 | 871 | 4 | US-09-105-058C-20 |
| 18 | 43 | 39.8 | 872 | 4 | US-09-177-650-2 |
| 19 | 43 | 39.8 | 930 | 4 | US-09-177-650-96 |
| 20 | 42 | 38.9 | 14 | 1 | US-08-315-461-6 |
| 21 | 42 | 38.9 | 14 | 1 | US-08-474-997-1 |
| 22 | 42 | 38.9 | 14 | 2 | US-08-796-598-19 |
| 23 | 42 | 38.9 | 14 | 2 | US-08-447-175A-19 |
| 24 | 42 | 38.9 | 14 | 3 | US-08-622-046B-19 |
| 25 | 42 | 38.9 | 14 | 4 | US-08-843-076D-44 |
| 26 | 42 | 38.9 | 292 | 1 | US-08-036-210-12 |
| 27 | 42 | 38.9 | 292 | 2 | US-08-449-609-12 |

| | | | | | | |
|----|------|------|------|---|----------------------|-------------------|
| 28 | 42 | 38.9 | 292 | 4 | US-09-361-096A-12 | Sequence 12, Appl |
| 29 | 42 | 38.9 | 522 | 4 | US-09-543-681A-5830 | Sequence 5830, Ap |
| 30 | 41.5 | 38.4 | 259 | 4 | US-09-252-991A-32902 | Sequence 32902, A |
| 31 | 41 | 38.0 | 93 | 4 | US-09-489-039A-13141 | Sequence 13141, A |
| 32 | 41 | 38.0 | 211 | 3 | US-09-075-194-1 | Sequence 1, Appli |
| 33 | 41 | 38.0 | 292 | 4 | US-09-543-681A-5819 | Sequence 5819, Ap |
| 34 | 41 | 38.0 | 347 | 4 | US-09-818-780-96 | Sequence 96, Appl |
| 35 | 41 | 38.0 | 485 | 3 | US-09-384-212-2 | Sequence 2, Appli |
| 36 | 41 | 38.0 | 490 | 1 | US-08-294-770A-2 | Sequence 2, Appli |
| 37 | 41 | 38.0 | 491 | 2 | US-08-448-735C-2 | Sequence 3, Appli |
| 38 | 41 | 38.0 | 497 | 4 | US-09-181-339-3 | Sequence 6265, Ap |
| 39 | 41 | 38.0 | 1014 | 4 | US-09-134-000C-6265 | Sequence 1, Appli |
| 40 | 40 | 37.0 | 13 | 1 | US-08-021-839A-1 | Sequence 6, Appli |
| 41 | 40 | 37.0 | 13 | 1 | US-08-218-608-6 | Sequence 12, Appl |
| 42 | 40 | 37.0 | 13 | 2 | US-08-792-553-12 | Sequence 46, Appl |
| 43 | 40 | 37.0 | 13 | 4 | US-09-129-192C-46 | Sequence 5, Appli |
| 44 | 40 | 37.0 | 14 | 2 | US-08-360-784B-5 | Sequence 5, Appli |
| 45 | 40 | 37.0 | 14 | 3 | US-09-054-308A-5 | Sequence 5, Appli |

ALIGNMENTS

RESULT 1
US-08-475-955-58
; Sequence 58, Application US/08475955
; Patent No. 6641813
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; TITLE OF INVENTION: METHODS AND REAGENTS FOR DIAGNOSIS OF
; NUMBER OF SEQUENCES: 218
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1250 West Peachtree Street
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,955
; FILING DATE: June 7, 1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/867,819
; FILING DATE: April 13, 1992
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/648,205
; FILING DATE: January 31, 1991
; APPLICATION DATA: 07/472,947
; FILING DATE: January 31, 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF114CIP(2)DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-873-8794
; TELEFAX: (404)-873-8795
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:

```
; NAME/KEY: Binding-site
; LOCATION: 6..15
; US-08-475-955-58
Query Match 47.2%; Score 51; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KARIHPFH 9
Db 9 KARIHPFH 17

RESULT 2
US-08-716-249-11
; Sequence 11, Application US/08716249
; Patent No. 6455244
; GENERAL INFORMATION:
; APPLICANT: Guichard, Gilles
; APPLICANT: Muller, Sylviane
; APPLICANT: Briand, Jean-Paul
; APPLICANT: Regemortel, Marc
; TITLE OF INVENTION: Retropeptides, Antibodies Thereto, and
; TITLE OF INVENTION: Uses Thereof for Vaccination and In Vitro Diagnosis
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spencer & Frank
; STREET: 1100 New York Avenue, Suite 300E
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/716,249
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/00292
; FILING DATE: 13-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Calvetti, Frederick F.
; REGISTRATION NUMBER: 28,557
; REFERENCE/DOCKET NUMBER: GROFO 7001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)414-4000
; TELEFAX: (202)414-4040
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-716-249-11
Query Match 47.2%; Score 51; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KARIHPFH 9
Db 13 KARIHPFH 21

RESULT 3
US-08-160-604-87
; Sequence 87, Application US/08160604
; Patent No. 6232522
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; APPLICANT: Scofield, R. H.
; TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOGY
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; US-08-160-604-87
Query Match 43.5%; Score 47; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KARIHPFH 8
Db 1 KARIHPFH 8

RESULT 4
US-08-160-604-78
; Sequence 78, Application US/08160604
; Patent No. 6232522
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; APPLICANT: Scofield, R. H.
; TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOGY
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; US-08-160-604-87
Query Match 43.5%; Score 47; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KARIHPFH 8
Db 1 KARIHPFH 8
```


Thu May 20 11:48:31 2004

us-10-005-684-6.ra1

Page 3

STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160,604
FILING DATE: 30-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/867,819
FILING DATE: 13-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/648,205
FILING DATE: 31-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,947
FILING DATE: 31-JAN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF114CIP(3)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-815-6508
TELEFAX: (404)-815-6555
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-160-604-78

Query Match 43.5%; Score 47; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KARIHPH 8
Db 9 KARIHPH 16

RESULT 5
US-08-033-797-2
; Sequence 2, Application US/08033797
; Patent No. 5583210
; GENERAL INFORMATION:
; APPLICANT: NEILL, John
; APPLICANT: PIERCE, Dorothy A.
; TITLE OF INVENTION: Methods and Compositions for Controlling
; TITLE OF INVENTION: Plant Development
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESS: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/160,604
; FILING DATE: 30-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/867,819
; FILING DATE: 13-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF114CIP(3)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-815-6508
; TELEFAX: (404)-815-6555
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-160-604-78

FILING DATE: 18-MAR-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 33229/168/PIHI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 220 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-033-797-2

Query Match 42.6%; Score 46; DB 1; Length 220;
Best Local Similarity 60.0%; Pred. No. 3.4;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KARIHPH 10
Db 88 RVRHPFHL 97

RESULT 6
US-08-472-265-2
; Sequence 2, Application US/08472265
; Patent No. 5728817
; GENERAL INFORMATION:
; APPLICANT: NEILL, John
; APPLICANT: PIERCE, Dorothy A.
; APPLICANT: CIGAN, Andrew M.
; TITLE OF INVENTION: Methods and Compositions for Controlling
; TITLE OF INVENTION: Plant Development
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESS: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,265
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/033,797
; FILING DATE: 18-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 33229/335/PIHI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 220 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-472-265-2

Query Match 42.6%; Score 46; DB 1; Length 220;
Best Local Similarity 60.0%; Pred. No. 3.4;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KARIHPHIL 10
: : : : :
Db 88 RRVHPHVL 97

RESULT 7

US-08-472-263-2
; Sequence 2, Application US/08472263
; Patent No. 5760190
; GENERAL INFORMATION:
; APPLICANT: NEILL, John
; APPLICANT: PIERCE, Dorothy A.
; APPLICANT: CIGAN, Andrew M.
; TITLE OF INVENTION: Methods and Compositions for Controlling
; TITLE OF INVENTION: Plant Development
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472.263
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/033,797
; FILING DATE: 18-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 33229/334/PIHI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 220 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-472-263-2

Query Match 42.6%; Score 46; DB 1; Length 220;
Best Local Similarity 60.0%; Pred. No. 3.4;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KARIHPHIL 10
: : : : :
Db 88 RRVHPHVL 97

RESULT 8

US-09-866-108A-3
; Sequence 3, Application US/09866108A
; Patent No. 6686188
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE

; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108A
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 15755
; SOFTWARE: Aeomica Sequence Listing Engine
; Patent No. 6686188
; SEQ ID NO 3
; LENGTH: 2568
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-108A-3

Query Match 40.7%; Score 44; DB 4; Length 2568;
Best Local Similarity 40.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KARIHPHILAEITYKIGH 20
: : : : :
Db 1242 RVQLAGPHLEALRLHRTGY 1261

RESULT 9

US-08-033-797-3
; Sequence 3, Application US/08033797
; Patent No. 5583210
; GENERAL INFORMATION:
; APPLICANT: NEILL, John
; APPLICANT: PIERCE, Dorothy A.
; APPLICANT: CIGAN, Andrew M.
; TITLE OF INVENTION: Methods and Compositions for Controlling
; TITLE OF INVENTION: Plant Development
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/033,797
; FILING DATE: 18-MAR-1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 33229/168/PIHI
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 214 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-033-797-3

Query Match 39.8%; Score 43; DB 1; Length 214;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPHIL 10
: : : : :
Db 88 RVLHPHVI 97

RESULT 10

US-08-472-265-3
Sequence 3, Application US/08472265
Patent No. 5728817

GENERAL INFORMATION:
APPLICANT: NEILL, John
APPLICANT: PIERCE, Dorothy A.
APPLICANT: CIGAN, Andrew M.
TITLE OF INVENTION: Methods and Compositions for Controlling
TITLE OF INVENTION: Plant Development
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,265
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/033,797
FILING DATE: 18-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 33229/335/PIHI
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 214 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-472-265-3

Query Match 39.8%; Score 43; DB 1; Length 214;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPHIL 10
: : : : :
Db 88 RVLHPHVI 97

RESULT 11

US-08-472-263-3

Sequence 3, Application US/08472263
Patent No. 5760190
GENERAL INFORMATION:
APPLICANT: NEILL, John
APPLICANT: PIERCE, Dorothy A.
APPLICANT: CIGAN, Andrew M.
TITLE OF INVENTION: Methods and Compositions for Controlling
TITLE OF INVENTION: Plant Development
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,263
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/033,797
FILING DATE: 18-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 33229/334/PIHI
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 214 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-472-263-3

Query Match 39.8%; Score 43; DB 1; Length 214;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHPHIL 10
: : : : :
Db 88 RVLHPHVI 97

RESULT 12

US-09-328-352-6571
Sequence 6571, Application US/09328352
Patent No. 6562958

GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 6571
LENGTH: 214
TYPE: PRT

ORGANISM: Acinetobacter baumannii
US-09-328-352-6571

Query Match 39.8%; Score 43; DB 4; Length 214;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RIHPFHILIA 12

```
Db      196 KVHPLHVLIA 205
      :||| |:|
RESULT 13
US-09-105-058C-23
; Sequence 23, Application US/09105058C
; Patent No. 6403360
; GENERAL INFORMATION:
; APPLICANT: Blamar, Michael A.
; APPLICANT: Dworetzky, Steven
; APPLICANT: Gribkoff, Valentin K.
; APPLICANT: Levesque, Paul C.
; APPLICANT: Little, Wayne A.
; APPLICANT: Neubauer, Michael G.
; APPLICANT: Yang, Wen-Pin
; TITLE OF INVENTION: KCNQ POTASSIUM CHANNELS AND METHODS OF MODULATING SAME
; FILE REFERENCE: 3053-4052
; CURRENT APPLICATION NUMBER: US/09/105,058C
; CURRENT FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US 60/055,599
; PRIOR FILING DATE: 1997-08-12
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 722
; TYPE: PRT
; ORGANISM: mouse
US-09-105-058C-23
Query Match      39.8%; Score 43; DB 4; Length 722;
Best Local Similarity 30.0%; Pred. No. 43;
Matches 6; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      1 KARIHPFHILIALEYTKGH 20
      | : | : | : |
Db      528 KESLRPYDVMVDIEQYSAGH 547

RESULT 14
US-09-177-650-89
; Sequence 89, Application US/09177650
; Patent No. 6413719
; GENERAL INFORMATION:
; APPLICANT: Leppert, Mark F.
; APPLICANT: Singh, Nanda
; APPLICANT: Charlier, Carole
; TITLE OF INVENTION: KCNQ2 AND KCNQ3 - POTASSIUM CHANNEL GENES WHICH ARE
; TITLE OF INVENTION: MUTATED IN BENIGN FAMILIAL NEONATAL CONVULSIONS (BFNC)
; TITLE OF INVENTION: AND OTHER EPILEPSIES
; FILE REFERENCE: 2323-134
; CURRENT APPLICATION NUMBER: US/09/177,650
; CURRENT FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 60/063,147
; EARLIER FILING DATE: 1997-10-24
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
; LENGTH: 757
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-177-650-89
Query Match      39.8%; Score 43; DB 4; Length 757;
Best Local Similarity 30.0%; Pred. No. 45;
Matches 6; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      1 KARIHPFHILIALEYTKGH 20
      | : | : | : |
Db      527 KESLRPYDVMVDIEQYSAGH 546

RESULT 15
US-09-813-148-4
; Sequence 4, Application US/09813148
; Patent No. 6617131
; GENERAL INFORMATION:
; APPLICANT: STEINMEYER, Klaus
; APPLICANT: LERCHE, Christian
; APPLICANT: SCHERER, Constanze
; APPLICANT: SEBOHM, Guiscard
; APPLICANT: BUSCH, Andreas E.
; TITLE OF INVENTION: POTASSIUM CHANNEL PROTEIN KCNQ5, A NEW TARGET FOR DISEASES OF CENTRAL NERVOUS SYSTEM AND CARDIOVASCULAR SYSTEM
; FILE REFERENCE: 38005-119
; CURRENT APPLICATION NUMBER: US/09/813,148
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: DE 100 13 732.6
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/194,041
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 844
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-813-148-4
Query Match      39.8%; Score 43; DB 4; Length 844;
Best Local Similarity 30.0%; Pred. No. 51;
Matches 6; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      1 KARIHPFHILIALEYTKGH 20
      | : | : | : |
Db      528 KESLRPYDVMVDIEQYSAGH 547

Search completed: May 19, 2004, 15:29:40
Job time : 17.5 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 15:28:52 ; Search time 59.5833 Seconds
(without alignments)
94.841 Million cell updates/sec

Title: US-10-005-684-6

Perfect score: 108

Sequence: 1 KARIHPFHLLIALEYKTH 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Geneseqp29Jan04:*
1: geneseqp1980s:*
2: geneseqp1980s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 108 | 100.0 | 20 | 7 ADB81379 | Adb81379 Lupus pep |
| 2 | 108 | 100.0 | 413 | 2 AAW03717 | Aaw03717 Human aut |
| 3 | 108 | 100.0 | 538 | 5 ABG32657 | Abg32657 Human adh |
| 4 | 108 | 100.0 | 552 | 4 AAB46827 | Aab46827 Human ant |
| 5 | 108 | 100.0 | 553 | 4 AAB46829 | Aab46829 Human ant |
| 6 | 108 | 100.0 | 553 | 4 AAB46828 | Aab46828 Human ant |
| 7 | 108 | 100.0 | 557 | 4 AAB94864 | Aab94864 Human pro |
| 8 | 51 | 47.2 | 17 | 2 AAR43468 | Aar43468 Ro/SSA ep |
| 9 | 49 | 45.4 | 83 | 6 ABP58906 | Abp58906 Human tra |
| 10 | 47 | 43.5 | 8 | 4 AAE08183 | Aae08183 Peptide # |
| 11 | 47 | 43.5 | 16 | 4 AAO12893 | Aao12893 Peptide # |
| 12 | 47 | 43.5 | 108 | 4 AAO12893 | Aao12893 Human pol |
| 13 | 46 | 42.6 | 183 | 3 AAG34111 | Aag34111 Zea mays |
| 14 | 46 | 42.6 | 220 | 2 AAR60770 | Aar60770 Maize OM |
| 15 | 46 | 42.6 | 220 | 3 AAG34110 | Aag34110 Zea mays |
| 16 | 46 | 42.6 | 240 | 3 AAG34109 | Aag34109 Zea mays |
| 17 | 46 | 42.6 | 877 | 5 ABB55496 | Abb55496 Lactococc |
| 18 | 45 | 41.7 | 74 | 3 AAG14412 | Aag14412 Arabidops |
| 19 | 45 | 41.7 | 111 | 3 AAG14411 | Aag14411 Arabidops |
| 20 | 45 | 41.7 | 111 | 3 AAG18444 | Aag18444 Arabidops |
| 21 | 45 | 41.7 | 126 | 3 AAG14410 | Aag14410 Arabidops |
| 22 | 45 | 41.7 | 148 | 3 AAG18443 | Aag18443 Arabidops |
| 23 | 45 | 41.7 | 166 | 3 AAG18442 | Aag18442 Arabidops |
| 24 | 45 | 41.7 | 184 | 3 AAG07200 | Aag07200 Arabidops |
| 25 | 45 | 41.7 | 184 | 3 AAG48968 | Aag48968 Arabidops |

ALIGNMENTS

RESULT 1

ID ADB81379 standard; peptide; 20 AA.

XX ADB81379;

XX 04-DEC-2003 (first entry)

XX Lupus peptide (20-mer) used in a test for antibodies against lupus.

XX immunoassay; autoimmune disease; autoantigen; ELISA;

XX enzyme-linked immunosorbant assay; saliva IgA; lupus; arthritis;

XX platelet glycoprotein; immune complex.

XX Synthetic.

XX US2003100035-A1.

XX 29-MAY-2003.

XX 08-NOV-2001; 2001US-00005684.

XX 08-NOV-2001; 2001US-00005684.

XX (VOJD/) VOJDANI A.

XX Vojdani A;

XX WPI; 2003-606630/57.

XX Saliva immunoassay for detection of antibodies for autoimmune disease,

XX e.g. lupus, in patient, by determining level of antibodies against

XX autoantigen, and comparing level of determined antibodies with normal

XX levels of antibodies.

XX Example 8; Page 7; 16pp; English.

XX This invention relates to a novel saliva immunoassay for diagnosing an

XX autoimmune disease in a patient. Specifically, the method determines the

XX level of antibodies that are present against the autoantigens for a

XX particular autoimmune disease and compares these results with normal

XX levels to determine the likelihood or severity of such a disease. The

XX test comprises a highly sensitive and accurate ELISA (enzyme-linked

XX immunosorbant assay) that measures saliva IgA specific antibody titres

XX against the purified antigens or a corresponding recombinant antigen or

XX synthetic peptide of an autoantigen, where the autoantigens are lupus

XX peptides, arthritis peptides, platelet glycoprotein or immune complexes.

XX As such, this single test can be used to accurately detect antibodies for

Aag07573 Arabidops
Aag34633 Arabidops
Aag07199 Arabidops
Aag48967 Arabidops
Aag07572 Arabidops
Aag07198 Arabidops
Aag48966 Arabidops
Aag34632 Arabidops
Aag07571 Arabidops
Aag34631 Arabidops
Aag14467 Novel hum
Abu02805 S. pneumo
Abu45768 Protein e
Abp27172 Streptoco
Aau37643 Streptoco
Aao01699 Human pol
Aag31268 Arabidops
Aag31267 Arabidops
Aag31266 Arabidops
Aae01905 Arabidops

CC diseases such as lupus or arthritis. Furthermore, it is useful to
 CC indicate ongoing pathology or to predict an early pathogenic reaction for
 CC autoimmune disease. This peptide sequence is the lupus peptide (20-mer)
 CC used in a test for lupus antibodies, in an exemplification of the
 CC invention.

XX Sequence 20 AA;
 SQ

Query Match 100.0%; Score 108; DB 7; Length 20;
 Best Local Similarity 100.0%; Pred. No. 5.3e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KARIHPHILIALEYTKGH 20

DB 1 KARIHPHILIALEYTKGH 20

RESULT 2

AAW03717 ID AAW03717 standard; protein; 413 AA.

AC AAW03717;

XX

DT 25-MAR-2003 (revised)

DT 12-MAR-1997 (first entry)

XX

DE Human autoantigen Ro(SS-A).

XX

KW Autoimmune disease; La autoantigen; Sjogren's syndrome;
 KW systemic lupus erythematosus; diagnosis.

XX

OS Homo sapiens.

PH Key

Protein Location/Qualifiers

1. .413

/label= Ro antigen

/note= "X's in the present sequence indicate amino acid
 positions given the designation End in the specification"

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

DB 150 KARIHPHILIALEYTKGH 169

RESULT 3

ABG32657

ID ABG32657 standard; protein; 538 AA.

XX

XX

XX

DT 15-NOV-2002 (first entry)

XX

DE Human adhesion molecule, termed P10155 (Ro60 or ADS5), protein.

XX

KW Human; adhesion molecule; ADS1; ADS2; ADS5; AAC74854.1; Ro60; AAC76768.1;

KW P10155; embryogenesis; apoptosis; homeostasis; phenotype; ligand;

KW diagnosis; vaccine; therapeutic; transgenic; knockout; atherosclerosis;

KW ischaemia; restenosis; reperfusion injury; sepsis; thrombosis; cancer;

KW haematological disease; leukaemia; blood clotting; thrombosis; cancer;

KW tumour; metastasis; inflammatory disease; rhinitis;

KW gastrointestinal disease; inflammatory bowel disease; ulcerative colitis;

KW Crohn's disease; respiratory disease; asthma; COPD;

KW chronic obstructive pulmonary disease; respiratory distress syndrome;

KW pulmonary fibrosis; immune disorder; autoimmune disease;

KW rheumatoid arthritis; transplant rejection; allergy; liver disease;

KW cirrhosis; endocrine disease; diabetes; bone disease; osteoporosis;

KW neurological disease; stroke; multiple sclerosis; spinal cord injury;

KW burn; wound healing; infection; cell-cell adhesion; gene therapy.

XX

XX

OS Homo sapiens.

XX

PH Key

Region Location/Qualifiers

373. .503

/note= "ADS5 adhesion molecule region"

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

CC diseases such as lupus or arthritis. Furthermore, it is useful to
 CC indicate ongoing pathology or to predict an early pathogenic reaction for
 CC autoimmune disease. This peptide sequence is the lupus peptide (20-mer)
 CC used in a test for lupus antibodies, in an exemplification of the
 CC invention.

XX Sequence 20 AA;
 SQ

Query Match 100.0%; Score 108; DB 2; Length 413;
 Best Local Similarity 100.0%; Pred. No. 1.5e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KARIHPHILIALEYTKGH 20

DB 1 KARIHPHILIALEYTKGH 20

RESULT 2

AAW03717 ID AAW03717 standard; protein; 413 AA.

AC AAW03717;

XX

DT 25-MAR-2003 (revised)

DT 12-MAR-1997 (first entry)

XX

DE Human autoantigen Ro(SS-A).

XX

KW Autoimmune disease; La autoantigen; Sjogren's syndrome;
 KW systemic lupus erythematosus; diagnosis.

XX

OS Homo sapiens.

PH Key

Protein Location/Qualifiers

1. .413

/label= Ro antigen

/note= "X's in the present sequence indicate amino acid
 positions given the designation End in the specification"

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

DB 150 KARIHPHILIALEYTKGH 169

RESULT 3

ABG32657

ID ABG32657 standard; protein; 538 AA.

XX

XX

XX

DT 15-NOV-2002 (first entry)

XX

DE Human adhesion molecule, termed P10155 (Ro60 or ADS5), protein.

XX

KW Human; adhesion molecule; ADS1; ADS2; ADS5; AAC74854.1; Ro60; AAC76768.1;

KW P10155; embryogenesis; apoptosis; homeostasis; phenotype; ligand;

KW diagnosis; vaccine; therapeutic; transgenic; knockout; atherosclerosis;

KW ischaemia; restenosis; reperfusion injury; sepsis; thrombosis; cancer;

KW haematological disease; leukaemia; blood clotting; thrombosis; cancer;

KW tumour; metastasis; inflammatory disease; rhinitis;

KW gastrointestinal disease; inflammatory bowel disease; ulcerative colitis;

KW Crohn's disease; respiratory disease; asthma; COPD;

KW chronic obstructive pulmonary disease; respiratory distress syndrome;

KW pulmonary fibrosis; immune disorder; autoimmune disease;

KW rheumatoid arthritis; transplant rejection; allergy; liver disease;

KW cirrhosis; endocrine disease; diabetes; bone disease; osteoporosis;

KW neurological disease; stroke; multiple sclerosis; spinal cord injury;

KW burn; wound healing; infection; cell-cell adhesion; gene therapy.

XX

XX

OS Homo sapiens.

XX

PH Key

Region Location/Qualifiers

373. .503

/note= "ADS5 adhesion molecule region"

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

DB 150 KARIHPHILIALEYTKGH 169

RESULT 3

ABG32657

ID ABG32657 standard; protein; 538 AA.

XX

XX

XX

DT 15-NOV-2002 (first entry)

XX

DE Human adhesion molecule, termed P10155 (Ro60 or ADS5), protein.

XX

KW Human; adhesion molecule; ADS1; ADS2; ADS5; AAC74854.1; Ro60; AAC76768.1;

KW P10155; embryogenesis; apoptosis; homeostasis; phenotype; ligand;

KW diagnosis; vaccine; therapeutic; transgenic; knockout; atherosclerosis;

KW ischaemia; restenosis; reperfusion injury; sepsis; thrombosis; cancer;

KW haematological disease; leukaemia; blood clotting; thrombosis; cancer;

KW tumour; metastasis; inflammatory disease; rhinitis;

KW gastrointestinal disease; inflammatory bowel disease; ulcerative colitis;

KW Crohn's disease; respiratory disease; asthma; COP

PT stroke).

XX Claim 1; Fig 23; 125pp; English.

PS

XX The invention discloses human adhesion molecule polypeptides (designated

CC ADSs) ADS1, ADS2 and ADS5 which have the NCBI Genbank accession numbers,

CC and also termed, AAC74854.1, AAC76768.1 and P10155, respectively.

CC Adhesion molecules are involved in a wide range of biological processes

CC from embryogenesis to apoptosis. They are essential to the structural

CC integrity and homeostatic functioning of most tissues and, therefore,

CC defects can play a role in many disease processes. Alteration of their

CC activity is a means to alter the disease phenotype. The polynucleotides

CC and polypeptides can be used to identify ligands that bind specifically

CC to, and which preferably inhibit, the adhesion molecule activity, to

CC identify compounds that either increase or decrease the level of

CC expression or activity, for diagnosing a disease in a patient, in a

CC pharmaceutical composition, as a vaccine, for monitoring the therapeutic

CC treatment of disease in a patient, for creating transgenic or knockout

CC animals, expressing higher, lower or absent levels of the ADS polypeptide

CC and for screening a compound effective for treating a disease. The ADS

CC polypeptides, nucleic acids, vectors, ligands, compounds or

CC pharmaceutical compositions are useful for treating or diagnosing

CC diseases including atherosclerosis, ischaemia, stenosis, reperfusion

CC injury, sepsis, haematological diseases (e.g. leukaemia), blood clotting

CC disorders (e.g. thrombosis), cancer (e.g. lung, prostate, breast,

CC colorectal or brain tumours), metastasis, inflammatory diseases (e.g.

CC rhinitis), gastrointestinal diseases (e.g. inflammatory bowel disease,

CC ulcerative colitis or Crohn's disease), respiratory diseases (e.g.

CC asthma, chronic obstructive pulmonary disease (COPD), respiratory

CC distress syndrome or pulmonary fibrosis), immune disorders (e.g.

CC autoimmune diseases, rheumatoid arthritis, transplant rejection or

CC allergy), liver diseases (e.g. cirrhosis), endocrine diseases (e.g.

CC diabetes), bone diseases (e.g. osteoporosis), neurological diseases (e.g.

CC stroke or multiple sclerosis), spinal cord injury, burns and wound

CC healing or infections (e.g. bacterial infections, particularly

CC Escherichia coli infection). The ADS polypeptide is particularly useful

CC as an adhesion molecule or for effecting cell-cell adhesion. The protein

CC sequence presented is the human adhesion molecule, termed P10155 (Ro60 or

XX ADS5)

XX

SQ Sequence 538 AA;

Query Match 100.0%; Score 108; DB 5; Length 538;

Best Local Similarity 100.0%; Pred. No. 2e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KARIHPHILIALETYKTH 20

Db 316 KARIHPHILIALETYKTH 335

RESULT 4

AAB46827

ID AAB46827 standard; protein; 552 AA.

AC AAB46827;

DT 26-APR-2001 (first entry)

DE Human antigen SSA60 M4-C6 protein.

XX Antigen; human; SSA60; ribonucleoprotein; autoimmune disease; diagnosis;

XX lupus erythematosus; Sjorgen's syndrome A.

XX Homo sapiens.

OS

XX DE19931380-A1.

FN 11-JAN-2001.

PD

XX 07-JUL-1999; 99DE-01031380.

PF

XX 07-JUL-1999; 99DE-01031380.

XX

PT

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Burckhardt J, Haass M, Lehmann H;

PI

XX WPI; 2001-148274/16.

DR

XX Recombinant production of ribonucleoprotein, useful for diagnosis of

PT autoimmune disease, by expressing, in prokaryotes, sequences encoding

XX both RNA and protein components.

XX

PS Example 1; Page 13-15; 36pp; German.

XX

CC This invention describes a novel method for the recombinant production of

CC ribonucleoprotein (I) which comprises producing a prokaryotic host cell

CC that contains at least one DNA (II) encoding a protein component of (I) and

CC at least one DNA (III) encoding a protein component of (I), expressing

CC both DNAs and then recovering (I). The invention also describes (1)

CC nucleic acid constructs containing (II) and (III); (2) recombinant

CC prokaryotic cell containing (II) and (III); (3) (I) produced by the new

CC method; and (4) SSA60 protein, designated M56, having a 553 amino acid

CC (aa) sequence, given in the specification, optionally in association with

CC RNA. (I) are used for diagnosis/prognosis of autoimmune diseases,

CC especially systemic lupus erythematosus or Sjorgen's syndrome A, by

CC complex formation with (I)-specific antibodies. (I) are produced simply,

CC inexpensively and on a large scale, in a functional (immunologically

CC active) form that provides a reliable diagnostic test. Complete (I) are

CC more selective and sensitive than the protein component used alone,

CC whether recombinant or isolated from bovine spleen

XX

SQ Sequence 552 AA;

Query Match 100.0%; Score 108; DB 4; Length 552;

Best Local Similarity 100.0%; Pred. No. 2.1e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KARIHPHILIALETYKTH 20

Db 328 KARIHPHILIALETYKTH 347

RESULT 5

AAB46829

ID AAB46829 standard; protein; 553 AA.

AC AAB46829;

DT 26-APR-2001 (first entry)

DE Human antigen construct pEQ30-HY3-SSA60M56 #4 protein.

XX Antigen; human; SSA60; ribonucleoprotein; autoimmune disease; diagnosis;

XX lupus erythematosus; Sjorgen's syndrome A.

XX Homo sapiens.

OS

XX DE19931380-A1.

FN 11-JAN-2001.

PD

XX 07-JUL-1999; 99DE-01031380.

PF

XX 07-JUL-1999; 99DE-01031380.

XX

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Burckhardt J, Haass M, Lehmann H;

PI

XX WPI; 2001-148274/16.

DR N-PSDB; AAF26157.

XX

PT Recombinant production of ribonucleoprotein, useful for diagnosis of

XX autoimmune disease, by expressing, in prokaryotes, sequences encoding

PT

PT both RNA and protein components.

PS Example 1; Page 22-24; 36pp; German.

XX
CC This invention describes a novel method for the recombinant production of ribonucleoprotein (I) which comprises producing a prokaryotic host cell that contains at least one DNA (II) encoding an RNA component of (I) and at least one DNA (III) encoding a protein component of (I), expressing both DNAs and then recovering (I). The invention also describes (1) nucleic acid constructs containing (II) and (III); (2) recombinant prokaryotic cell containing (II) and (III); (3) (I) produced by the new method; and (4) SSA60 protein, designated M56, having a 553 amino acid (aa) sequence, given in the specification, optionally in association with RNA. (I) are used for diagnosis/prognosis of autoimmune diseases, especially systemic lupus erythematosus or Sjorgen's syndrome A, by complex formation with (I)-specific antibodies. (I) are produced simply, inexpensively and on a large scale, in a functional (immunologically active) form that provides a reliable diagnostic test. Complete (I) are more selective and sensitive than the protein component used alone, whether recombinant or isolated from bovine spleen

XX Sequence 553 AA;

Query Match 100.0%; Score 108; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KARIHPFHILIALETYKTH 20
| | | | | | | | | | | | | | | | | | | | | |
DB 331 KARIHPFHILIALETYKTH 350

RESULT 6

ID AAB46828
ID AAB46828 standard; protein; 553 AA.

XX AAB46828;

DT 26-APR-2001 (first entry)

DE Human antigen SSA60 M56 protein.

XX Antigen; human; SSA60; ribonucleoprotein; autoimmune disease; diagnosis;
XX lupus erythematosus; Sjorgen's syndrome A.

XX Homo sapiens.

XX DE19931380-A1.

XX 11-JAN-2001.

XX 07-JUL-1999; 99DE-01031380.

XX 07-JUL-1999; 99DE-01031380.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Burckhardt J, Haass M, Lehmann H;

XX WPI; 2001-148274/16.

XX Recombinant production of ribonucleoprotein, useful for diagnosis of autoimmune disease, by expressing, in prokaryotes, sequences encoding both RNA and protein components.

PS Claim 15; Page 15-18; 36pp; German.

XX This invention describes a novel method for the recombinant production of ribonucleoprotein (I) which comprises producing a prokaryotic host cell that contains at least one DNA (II) encoding an RNA component of (I) and at least one DNA (III) encoding a protein component of (I), expressing both DNAs and then recovering (I). The invention also describes (1) nucleic acid constructs containing (II) and (III); (2) recombinant

CC prokaryotic cell containing (II) and (III); (3) (I) produced by the new method; and (4) SSA60 protein, designated M56, having a 553 amino acid (aa) sequence, given in the specification, optionally in association with RNA. (I) are used for diagnosis/prognosis of autoimmune diseases, especially systemic lupus erythematosus or Sjorgen's syndrome A, by complex formation with (I)-specific antibodies. (I) are produced simply, inexpensively and on a large scale, in a functional (immunologically active) form that provides a reliable diagnostic test. Complete (I) are more selective and sensitive than the protein component used alone, whether recombinant or isolated from bovine spleen

XX Sequence 553 AA;

Query Match 100.0%; Score 108; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KARIHPFHILIALETYKTH 20
| | | | | | | | | | | | | | | | | | | | | |
DB 331 KARIHPFHILIALETYKTH 350

RESULT 7

AAB94864

ID AAB94864 standard; protein; 557 AA.

XX AAB94864;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:16064.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Iehii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

PS Claim 8; SEQ ID NO 16064; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

XX Sequence 557 AA;
 SQ

Query Match 100.0%; Score 108; DB 4; Length 557;
 Best Local Similarity 100.0%; Pred. No. 2.1e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KARIHPHILIALETYKTH 20
 |||||
 DB 335 KARIHPHILIALETYKTH 354

RESULT 8

AA43468
 ID AAR43468 standard; peptide; 17 AA.

XX AC AAR43468;

XX 25-MAR-2003 (revised)

DT 12-MAY-1994 (first entry)

XX Ro/SSA epitope 308.

DE Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

XX Homo sapiens.

OS WO9321223-A1.

PN 28-OCT-1993.

XX 13-APR-1993; 93WO-US003484.

PF 13-APR-1992; 92US-00867819.

PR (OKLA) UNIV OKLAHOMA STATE.

PA Harley JB;

PI WPI; 1993-351658/44.

XX New linear epitope(s) for human auto-antibodies - from the Ro/SSA, La/SSB
 PT and Sm B/B' antigens and ribonucleoprotein, used for diagnosing and
 PT treating auto-immune disorders e.g. systemic lupus erythematosus.

XX Claim 1; Page 31; 43pp; English.

XX The sequences given in AAR43391-562 are linear epitopes which are derived
 CC from the 60 kD Ro/SSA peptide, the La/SSB autoantigen, the 70 kD nuclear
 CC ribonucleoprotein (nRNP) and the Sm B/B' polypeptide. These antigens are
 CC common in systemic lupus erythematosus (SLE) and closely related
 CC disorders. The Ro/SSA family of proteins has been shown to have several
 CC molecular forms which are defined by the molecular weight of the antigen
 CC identified. The major form has a molecular weight of 60 kD and two
 CC additional forms have molecular weights of 52 and 54 kD. La/SSB is also a
 CC member of this group of autoantibodies and binds small RNAs with a
 CC polyuridine terminus. La/SSB is bound by a third of the anti-Ro/SSA
 CC precipitin positive sera. La/SSB has been shown to be a 46-50 kD
 CC monomeric phosphoprotein which associates with RNA polymerase III

CC transcripts. Anti-Sm antibodies precipitate snRNPs containing the U1, U2,
 CC U4/U6 and U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD), E/F
 CC (11 kD doublet) and G (less than 10 kD). These epitopes may be used for
 CC preventing, treating or screening autoimmune disorders, especially SLE or
 CC Sjogrens syndrome (SS). They bind to a human autoantibody and may
 CC therefore be used as vaccines. (Updated on 25-MAR-2003 to correct PN
 CC field.)

XX Sequence 17 AA;

Query Match 47.2%; Score 51; DB 2; Length 17;
 Best Local Similarity 100.0%; Pred. No. 0.41;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KARIHPFHI 9
 |||||
 DB 9 KARIHPFHI 17

RESULT 9

ABP58906
 ID ABP58906 standard; protein; 83 AA.

XX AC ABP58906;

XX 10-APR-2003 (first entry)

XX Human transcription termination factor binding protein 9.13.

DE Human; transcription termination factor binding protein 9.13;
 KW recombinant production; gene therapy; cancer; tumour; HIV infection;
 KW human immunodeficiency virus; cytostatic.

XX Homo sapiens.

XX CN1363566-A.

XX 14-AUG-2002.

XX 05-JAN-2001; 2001CN-00105036.

XX 05-JAN-2001; 2001CN-00105036.

XX (BODE-) BODE GENE DEV CO LTD SHANGHAI.

XX Mao Y, Xie Y;

XX WPI; 2003-000307/01.

XX N-PSDB; AB257684.

XX Polypeptide-human transcription termination factor bindin 9.13 and
 PT polynucleotide for coding it.

XX Claim 1; Page 26 (Disclosure); 32pp; Chinese.

XX The invention relates to human transcription termination factor binding
 CC protein 9.13 (ABP58906) and nucleic acids encoding it (AB257684). The
 CC protein has a molecular weight of 9.13 kD. The invention also relates to
 CC a method for the recombinant production of the protein, an antagonist of
 CC the protein, and the use of the protein, gene and antagonist in
 CC therapeutic applications. Transcription termination factor binding
 CC protein 9.13 can be used in the treatment of a variety of diseases such
 CC as cancer and HIV (human immunodeficiency virus) infection. The present
 CC sequence represents human transcription termination factor binding
 CC protein 9.13

XX Sequence 83 AA;

Query Match 45.4%; Score 49; DB 6; Length 83;
 Best Local Similarity 56.2%; Pred. No. 4.9;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Db 9 KARIHPFH 16

RESULT 12
AAO12893
ID AAO12893 standard; protein; 108 AA.
XX
XX AAO12893;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human polypeptide SEQ ID NO 26785.
DE
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
OS
XX WO200164835-A2.
PN
XX
XX 07-SEP-2001.
PD
XX
XX 26-FEB-2001; 2001WO-US004927.
PF
XX
XX 28-FEB-2000; 2000US-00515126.
PR
XX 18-MAY-2000; 2000US-00577409.
PR
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI: 2001-514838/56.
DR
XX N-PSDB; AAI92824.
DR
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
PT
XX
XX Claim 20; SEQ ID NO 26785; 1399pp + Sequence Listing; English.
PS
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 108 AA;
SQ

Query Match 43.5%; Score 47; DB 4; Length 108;
Best Local Similarity 64.3%; Pred. No. 13;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 PPHILIALEYKRG 19
| | | | |
Db 70 PPHIIKLSKYKRG 83
| | | | |

RESULT 13
AAG34111
ID AAG34111 standard; protein; 183 AA.
XX
XX AAG34111;
AC
XX
XX 18-OCT-2000 (first entry)
DT

XX DE Zea mays protein fragment SEQ ID NO: 41455.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
XX
XX Zea mays subsp. mays.
OS
XX EPI033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-00301439.
PF
XX
XX 25-FEB-1999; 99US-0121825P.
PR
XX 05-MAR-1999; 99US-0123180P.
PR
XX 09-MAR-1999; 99US-0123548P.
PR
XX 23-MAR-1999; 99US-0125788P.
PR
XX 25-MAR-1999; 99US-0126264P.
PR
XX 29-MAR-1999; 99US-0126785P.
PR
XX 01-APR-1999; 99US-0127462P.
PR
XX 06-APR-1999; 99US-0128234P.
PR
XX 08-APR-1999; 99US-0128714P.
PR
XX 16-APR-1999; 99US-0129845P.
PR
XX 19-APR-1999; 99US-0130077P.
PR
XX 21-APR-1999; 99US-0130449P.
PR
XX 23-APR-1999; 99US-0130510P.
PR
XX 28-APR-1999; 99US-0130891P.
PR
XX 30-APR-1999; 99US-0131449P.
PR
XX 30-APR-1999; 99US-0132048P.
PR
XX 30-APR-1999; 99US-0132407P.
PR
XX 04-MAY-1999; 99US-0132484P.
PR
XX 05-MAY-1999; 99US-0132485P.
PR
XX 06-MAY-1999; 99US-0132486P.
PR
XX 06-MAY-1999; 99US-0132487P.
PR
XX 07-MAY-1999; 99US-0132863P.
PR
XX 11-MAY-1999; 99US-0134256P.
PR
XX 14-MAY-1999; 99US-0134218P.
PR
XX 14-MAY-1999; 99US-0134219P.
PR
XX 14-MAY-1999; 99US-0134221P.
PR
XX 14-MAY-1999; 99US-0134370P.
PR
XX 18-MAY-1999; 99US-0134768P.
PR
XX 19-MAY-1999; 99US-0134941P.
PR
XX 20-MAY-1999; 99US-0135124P.
PR
XX 21-MAY-1999; 99US-0135353P.
PR
XX 24-MAY-1999; 99US-0135629P.
PR
XX 25-MAY-1999; 99US-0136021P.
PR
XX 27-MAY-1999; 99US-0136392P.
PR
XX 28-MAY-1999; 99US-0136782P.
PR
XX 01-JUN-1999; 99US-0137222P.
PR
XX 03-JUN-1999; 99US-0137528P.
PR
XX 04-JUN-1999; 99US-0137502P.
PR
XX 07-JUN-1999; 99US-0137724P.
PR
XX 08-JUN-1999; 99US-0138094P.
PR
XX 10-JUN-1999; 99US-0138540P.
PR
XX 10-JUN-1999; 99US-0138847P.
PR
XX 14-JUN-1999; 99US-0139119P.
PR
XX 16-JUN-1999; 99US-0139452P.
PR
XX 16-JUN-1999; 99US-0139453P.
PR
XX 17-JUN-1999; 99US-0139492P.
PR
XX 18-JUN-1999; 99US-0139454P.
PR
XX 18-JUN-1999; 99US-0139455P.
PR
XX 18-JUN-1999; 99US-0139456P.
PR
XX 18-JUN-1999; 99US-0139457P.
PR
XX 18-JUN-1999; 99US-0139458P.
PR
XX 18-JUN-1999; 99US-0139459P.
PR
XX 18-JUN-1999; 99US-0139460P.
PR
XX 18-JUN-1999; 99US-0139461P.
PR
XX 18-JUN-1999; 99US-0139462P.
PR
XX 18-JUN-1999; 99US-0139463P.
PR
XX 18-JUN-1999; 99US-0139750P.
PR
XX 18-JUN-1999; 99US-0139763P.

PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140833P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142085P.
PR 06-JUL-1999; 99US-0142330P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144844P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145087P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145132P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147132P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147433P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.

PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 42.6%; Score 46; DB 3; Length 183;
Best Local Similarity 60.0%; Pred. No. 34;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KARIHPHIL 10
Db 51 RVRVHPFVL 60

RESULT 14

AAR60770
ID AAR60770 standard; protein; 220 AA.

XX AC AAR60770;

XX DT 25-MAR-2003 (revised)

XX DT 04-JUN-1995 (first entry)

XX DE Maize QM homolog from clone 10-15.

XX KW Plant; transformation; sterile plants; hybrid seed.

XX OS Zea mays.

```

XX WO9421804-A1.
XX
XX 29-SEP-1994.
XX
XX 09-MAR-1994; 94WO-US002161.
XX
XX 18-MAR-1993; 93US-00033797.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Neill JD, Pierce DA, Cigan AM;
XX
XX WPI; 1994-317024/39.
XX
XX N-PSDB; AAQ73666.
XX
XX Polynucleotide(s) encoding a QM-type protein in plants - used to
XX transform plants to alter normal development partic. for producing male
XX sterile plants.
XX
XX Claim 9; Fig 2; 80pp; English.
XX
XX The sequence is that of a developmental protein in a plant which is
XX encoded by a DNA sequence capable of hybridising to a mammalian QM gene.
XX The DNA can be used to transform plants to alter normal development and
XX to produce male sterile plants, e.g. for the prodn. of hybrid seeds.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 220 AA;
XX
XX Query Match 42.6%; Score 46; DB 2; Length 220;
XX Best Local Similarity 60.0%; Pred. No. 42;
XX Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 KARIHPFHIL 10
XX Db 88 RRVHPFHL 97
XX
XX RESULT 15
XX AAG34110
XX ID AAG34110 standard; protein; 220 AA.
XX
XX AC AAG34110;
XX
XX DT 18-OCT-2000 (first entry)
XX
XX DE Zea mays protein fragment SEQ ID NO: 41454.
XX
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
XX
XX OS Zea mays subsp. mays.
XX
XX PN EF1033405-A2.
XX
XX PD 06-SEP-2000.
XX
XX PF 25-FEB-2000; 2000EP-00301439.
XX
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0129845P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.
XX
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131443P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140695P.
XX 28-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
XX 01-JUL-1999; 99US-0141842P.
XX 01-JUL-1999; 99US-0142154P.
XX 02-JUL-1999; 99US-0142055P.
XX 06-JUL-1999; 99US-0142390P.
XX 08-JUL-1999; 99US-0142803P.
XX 09-JUL-1999; 99US-0142920P.
XX 12-JUL-1999; 99US-0142977P.
XX 13-JUL-1999; 99US-0143542P.
XX 14-JUL-1999; 99US-0143624P.
XX 15-JUL-1999; 99US-0144005P.
XX 16-JUL-1999; 99US-0144085P.
XX 16-JUL-1999; 99US-0144086P.
XX 19-JUL-1999; 99US-0144325P.
XX 19-JUL-1999; 99US-0144331P.
XX 19-JUL-1999; 99US-0144332P.
XX 19-JUL-1999; 99US-0144333P.
XX 19-JUL-1999; 99US-0144334P.
XX 19-JUL-1999; 99US-0144335P.
XX 20-JUL-1999; 99US-0144352P.

```

PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147182P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148664P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.

PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160800P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 42.6%; Score 46; DB 3; Length 220;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KARIHFFHIL 10
Db : |||||:
88 RVRVHPFVL 97

Search completed: May 19, 2004, 15:43:56
Job time : 61.5833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 15:38:37 ; Search time 20.8333 Seconds
(without alignments)
151.449 Million cell updates/sec

Title: US-10-005-684-5
Perfect score: 60
Sequence: 1 SWEYSVWLSN 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriopl:*
- 17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|---------------------|
| 1 | 43 | 71.7 | 126 | Q9P862 | Q9P862 pyrenophora |
| 2 | 42 | 70.0 | 380 | Q93909 | Q93909 pseudomonas |
| 3 | 41 | 68.3 | 283 | Q9WIK0 | Q9WIK0 faba bean n |
| 4 | 41 | 68.3 | 283 | Q91250 | Q91250 faba bean n |
| 5 | 41 | 68.3 | 319 | Q9XW52 | Q9XW52 caenorhabdi |
| 6 | 41 | 68.3 | 2799 | Q812F6 | Q812F6 homo sapien |
| 7 | 40 | 66.7 | 283 | Q92034 | Q92034 milk vetch |
| 8 | 40 | 66.7 | 391 | Q88676 | Q88676 mus musculu |
| 9 | 40 | 66.7 | 502 | Q7UHL2 | Q7UHL2 rhodospirell |
| 10 | 40 | 66.7 | 539 | Q9JW85 | Q9JW85 neisseria m |
| 11 | 39 | 65.0 | 234 | Q86B64 | Q86B64 drosophila |
| 12 | 39 | 65.0 | 262 | Q46331 | Q46331 gracillariop |
| 13 | 39 | 65.0 | 277 | Q8DUU7 | Q8DUU7 streptococc |
| 14 | 39 | 65.0 | 300 | Q9J4N2 | Q9J4N2 human immun |
| 15 | 39 | 65.0 | 406 | P95129 | P95129 mycobacteri |
| 16 | 39 | 65.0 | 406 | Q7TXJ3 | Q7TXJ3 mycobacteri |

| | | | | | | |
|----|----|------|-----|----|--------|--------------------|
| 17 | 39 | 65.0 | 412 | 16 | Q9CNQ5 | Q9CNQ5 pasteurella |
| 18 | 39 | 65.0 | 450 | 16 | Q827K9 | Q827K9 streptomyce |
| 19 | 39 | 65.0 | 479 | 16 | Q9RKZ0 | Q9RKZ0 streptomyce |
| 20 | 39 | 65.0 | 808 | 10 | Q94GT7 | Q94GT7 oryza sativ |
| 21 | 39 | 65.0 | 808 | 10 | Q7XGN0 | Q7XGN0 oryza sativ |
| 22 | 39 | 65.0 | 820 | 10 | Q9FY05 | Q9FY05 populus tre |
| 23 | 39 | 65.0 | 938 | 10 | Q8SP55 | Q8SP55 oryza sativ |
| 24 | 38 | 63.3 | 77 | 17 | Q97VP9 | Q97VP9 sulfolobus |
| 25 | 38 | 63.3 | 135 | 17 | Q97X99 | Q97X99 sulfolobus |
| 26 | 38 | 63.3 | 148 | 17 | Q97XA2 | Q97XA2 sulfolobus |
| 27 | 38 | 63.3 | 185 | 16 | Q8YRG7 | Q8YRG7 anabaena sp |
| 28 | 38 | 63.3 | 212 | 16 | Q92EY4 | Q92EY4 listeria in |
| 29 | 38 | 63.3 | 212 | 16 | Q8YA63 | Q8YA63 listeria mo |
| 30 | 38 | 63.3 | 251 | 17 | Q97TW7 | Q97TW7 sulfolobus |
| 31 | 38 | 63.3 | 273 | 1 | Q9UWN3 | Q9UWN3 sulfolobus |
| 32 | 38 | 63.3 | 277 | 16 | Q8ED54 | Q8ED54 pseudomonas |
| 33 | 38 | 63.3 | 288 | 17 | Q97XC2 | Q97XC2 sulfolobus |
| 34 | 38 | 63.3 | 310 | 17 | Q97XK0 | Q97XK0 sulfolobus |
| 35 | 38 | 63.3 | 332 | 17 | Q97U87 | Q97U87 sulfolobus |
| 36 | 38 | 63.3 | 350 | 17 | Q97YS0 | Q97YS0 sulfolobus |
| 37 | 38 | 63.3 | 363 | 17 | Q9UWS3 | Q9UWS3 sulfolobus |
| 38 | 38 | 63.3 | 375 | 17 | Q97XG5 | Q97XG5 sulfolobus |
| 39 | 38 | 63.3 | 388 | 17 | Q97Y87 | Q97Y87 sulfolobus |
| 40 | 38 | 63.3 | 413 | 5 | Q9VKR5 | Q9VKR5 drosophila |
| 41 | 38 | 63.3 | 547 | 16 | Q8X4N4 | Q8X4N4 escherichia |
| 42 | 38 | 63.3 | 547 | 16 | Q8FKL3 | Q8FKL3 escherichia |
| 43 | 38 | 63.3 | 639 | 16 | Q8UHH1 | Q8UHH1 agrobacteri |
| 44 | 38 | 63.3 | 654 | 16 | Q8UBG0 | Q8UBG0 agrobacteri |
| 45 | 38 | 63.3 | 714 | 16 | Q8EMD1 | Q8EMD1 oceanobacil |

ALIGNMENTS

RESULT 1
Q9P862 ID Q9P862 PRELIMINARY; PRT; 126 AA.
AC Q9P862;
DT 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
DE Putative pathogenesis related protein.
GN PIG14.
OS Pyrenophora graminea (Barley leaf stripe).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
OC Pleosporales; Pleosporaceae; Pyrenophora.
OX NCBI_TaxID=5028;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dg2; TISSUE=Mycelium;
RA Vale G.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Dg2; TISSUE=Mycelium;
RA Bellini L.;
RT "Identification of Pyrenophora graminea induced genes during the
RT interaction with barley (Hordeum vulgare L.).";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ277800; CAB90218.1; -.
SQ SEQUENCE 126 AA; 14260 MW; 0494C2949575439A CRC64;

Query Match 71.7%; Score 43; DB 3; Length 126;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 WEYSVWLSN 10
Db 90 WRYSEWISN 98

RESULT 2
Q93909

```
ID Q939Q9 PRELIMINARY; PRT; 380 AA.
AC Q939Q9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 25, Last annotation update)
DE PHB depolymerase Phaz7 (Fragment).
GN PHAZ7.
OS Pseudomonas lemoignei.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Oxalobacteraceae; Paucimonas.
OX NCBI_TaxID=29443;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21453301; PubMed=11457823;
RA Handrick R., Reinhardt S., Focarete M.L., Scandola M., Adamus G.,
RA Kowalczyk M., Jendrosseck D.;
RT "A New Type of Thermoalkalophilic Hydrolase of Paucimonas lemoignei
RT with High Specificity for Amorphous Polyesters of Short Chain-length
RT Hydroxyalkanoic Acids.";
RL J. Biol. Chem. 276:36215-36224 (2001).
DR EMBL; AY026355; AAK07742.1; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0016298; F:lipase activity; IEA.
DR GO; GO:0016042; P:lipid catabolism; IEA.
DR InterPro; IPR002918; Lipase 2.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF01674; Lipase 2; 1.
FT NON TER 1
SQ SEQUENCE 380 AA; 40020 MW; 633E99A1391EEA CRC64;

Query Match 70.0%; Score 42; DB 2; Length 380;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 WEYSVWLSN 10
Db 240 WYGVVWSN 248

RESULT 3
Q9WIKO PRELIMINARY; PRT; 283 AA.
ID Q9WIKO
AC Q9WIKO;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Rep protein.
GN C7-EG.
OS Faba bean necrotic yellows virus.
OC Viruses; ssDNA viruses; Nanovirus.
OX NCBI_TaxID=59817;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Egyptian EV1-93;
MEDLINE=99094637; PubMed=9880028;
RA Katul L., Timchenko T., Gronenborn B., Vetten H.J.;
RT "Ten distinct circular ssDNA components, four of which encode putative
RT replication-associated proteins, are associated with the faba bean
RT necrotic yellows virus genome.";
RL J. Gen. Virol. 79:3101-3109 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Egyptian EV1-93;
MEDLINE=20027244; PubMed=10559333;
RA Timchenko T., de Kouchkovsky F., Katul L., David C., Vetten H.J.,
RA Gronenborn B.;
RT "A single Rep protein initiates replication of multiple genome
RT components of faba bean necrotic yellows virus, a single-stranded DNA
RT virus of plants.";
RL J. Virol. 73:10173-10182 (1999).
DR EMBL; AJ132185; CAB44025.1; -.
DR InterPro; IPR003365; Viral_rep.
DR Pfam; PF02407; Viral_rep; 1.

SQ SEQUENCE 283 AA; 33092 MW; EE49FFB3D450C060 CRC64;

Query Match 68.3%; Score 41; DB 12; Length 283;
Best Local Similarity 55.6%; Pred. No. 62;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WEYSVWLSN 10
Db 91 WEYGTWISS 99

RESULT 4
Q91250 PRELIMINARY; PRT; 283 AA.
ID Q91250
AC Q91250;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Component 7.
GN C7.
OS Faba bean necrotic yellows virus.
OC Viruses; ssDNA viruses; Nanovirus.
OX NCBI_TaxID=59817;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SV292-88;
MEDLINE=99094637; PubMed=9880028;
RA Katul L., Timchenko T., Gronenborn B., Vetten H.J.;
RT "Ten distinct circular ssDNA components, four of which encode putative
RT replication-associated proteins, are associated with the faba bean
RT necrotic yellows virus genome.";
RL J. Gen. Virol. 79:3101-3109 (1998).
DR EMBL; AJ005964; CAA06787.1; -.
DR InterPro; IPR003365; Viral_rep.
DR Pfam; PF02407; Viral_rep; 1.
SQ SEQUENCE 283 AA; 33150 MW; 0706D6ABB46E4067 CRC64;

Query Match 68.3%; Score 41; DB 12; Length 283;
Best Local Similarity 55.6%; Pred. No. 62;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WEYSVWLSN 10
Db 91 WEYGTWISS 99

RESULT 5
Q9XW52 PRELIMINARY; PRT; 319 AA.
ID Q9XW52
AC Q9XW52;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Y53C10A.5 protein.
GN Y53C10A.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA White S.;
RA Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99089613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018 (1998).
DR EMBL; AL033536; CAA22139.1; -.
DR PIR; T27118; T27118.
DR WormPep; Y53C10A.5; CE22373.
```


SQ SEQUENCE 319 AA; 36570 MW; 585814E71F55576A CRC64;
Query Match 68.3%; Score 41; DB 5; Length 319;
Best Local Similarity 66.7%; Pred. No. 70;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 SWEYSVWLS 9
DB 117 SWEYSVWLS 125
RESULT 6
Q81ZF6 PRELIMINARY; PRT; 2799 AA.
AC Q81ZF6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE G-protein coupled receptor GPR112.
GN GPR112.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fredriksson R., Lagerstrom M.C., Hoglund P., Schioth H.B.;
RT "New human G-protein coupled receptors with long N-terminals
containing GPS domains and Ser/Thr rich regions.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; A140954; AAN4668.1; -.
DR Genew; HGNC:18992; GPR112.
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:004930; F-G-protein coupled receptor activity; IEA.
DR GO; GO:0004872; F-receptor activity; IEA.
DR GO; GO:0007218; P-neuropeptide signaling pathway; IEA.
DR InterPro; IPR008985; ConA like lec_gl.
DR InterPro; IPR001623; DnaJ N.
DR InterPro; IPR000832; GPCR secretin.
DR InterPro; IPR000203; PKD_Cys_rich.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF01825; GPS; 1.
DR PRINTS; PR00249; GPCRSECRETIN.
DR SMART; SM00303; GPS; 1.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS0221; GPS; 1.
DR PROSITE; PS50261; G_PROTEIN_RECPT_F2_4; 1.
KW Receptor.
SQ SEQUENCE 2799 AA; 300964 MW; D3A31C9702A87B50 CRC64;
Query Match 68.3%; Score 41; DB 4; Length 2799;
Best Local Similarity 70.0%; Pred. No. 6.5e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 SWEYSVWLS 10
DB 144 SWEEDVWLVN 153
RESULT 7
Q92034 PRELIMINARY; PRT; 283 AA.
AC Q92034;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Replication initiation protein.
GN REP.
OS Milk vetch dwarf virus.
OC Viruses; ssDNA viruses; Nanovirus.
OX NCBI_TaxID=67585;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=99094638; PubMed=9880029;
RA Sano Y., Wada M., Hasimoto Y., Matsumoto T., Kojima M.;
RT "Sequences of ten circular ssDNA components associated with the milk
vetch dwarf virus genome.";
RL J. Gen. Virol. 79:3111-3118(1998).
DR EMBL; AB009047; BAA34048.1; -.
DR InterPro; IPR003365; Viral_rep.
DR Pfam; PF02407; Viral Rep; 1.
SQ SEQUENCE 283 AA; 33383 MW; 284EA4E36AE98078 CRC64;
Query Match 66.7%; Score 40; DB 12; Length 283;
Best Local Similarity 62.5%; Pred. No. 90;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 WEYSVWLS 9
DB 91 WEYGTWIS 98
RESULT 8
O88676 PRELIMINARY; PRT; 391 AA.
AC O88676;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CAMP metalloproteinase.
GN MMP23.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BALB/C; TISSUE=Spleen;
RX MEDLINE=99402935; PubMed=10471791;
RA Pei D.;
RT "CA-MMP: a matrix metalloproteinase with a novel cysteine array, but
without the classic cysteine switch(1).";
RL FEBS Lett. 457:262-270(1999).
DR EMBL; AF085742; AAC34886.1; -.
DR HSSP; P03956; ICGL.
DR MEROPS; M10.022; -.
DR MGD; MGI:1347361; Mmp23.
DR GO; GO:0005578; C-extracellular matrix; IEA.
DR GO; GO:0004222; F-metalloendopeptidase activity; IEA.
DR GO; GO:0005215; F-transporter activity; IEA.
DR GO; GO:0008270; F-zinc ion binding; IEA.
DR GO; GO:0008508; F-proteolysis and peptidolysis; IEA.
DR GO; GO:0006810; P-transport; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR000566; Lipocln_cytFABP.
DR InterPro; IPR006026; Peptidase_M.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR003582; ShKT.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00413; Peptidase_M10; 1.
DR Pfam; PF01549; SHK; 1.
DR PRINTS; PR00138; MATRXIN.
DR SMART; SM00409; IG; 1.
DR SMART; SM00254; ShKT; 1.
DR SMART; SM00235; ZmMC; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS00213; LIFOCALIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metal-binding; Metalloprotease; Protease; Zinc.
SQ SEQUENCE 391 AA; 44451 MW; 8C9675020F02F632 CRC64;
Query Match 66.7%; Score 40; DB 11; Length 391;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SWEYSVWLSN 10
 DB 196 SWKGVWLTN 205

RESULT 9
 Q7UHL2 PRELIMINARY; PRT; 502 AA.

AC Q7UHL2
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN R81326.

OS Rhodopirellula baltica.
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
 OC Planctomycetaceae; Pirellula.
 OX NCBI_TaxID=117;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=1;
 RX MEDLINE=22735913; PubMed=12835416;
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
 RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
 RA Schlesner H., Amann R., Reinhardt R.;
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.
 RT strain 1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
 DR ENBL; BX294156; CAD77958.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 502 AA; 55574 MW; D6015E54AE5044C4 CRC64;

Query Match 56.7%; Score 40; DB 16; Length 502;
 Best Local Similarity 75.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 WEYSVWLS 9
 DB 425 WGVSVWLA 432

RESULT 10
 Q9JVB5 PRELIMINARY; PRT; 539 AA.

AC Q9JVB5
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical protein NMA0954.
 GN NMA0954.
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=65699;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
 RX MEDLINE=20222556; PubMed=10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holtrope S.,
 RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrall B.G.;
 RT "Complete DNA sequence of a serogroup A strain of Neisseria
 RT meningitidis Z2491.";
 RL Nature 404:502-506 (2000).
 CC -1- COFACTOR: FAD (BY SIMILARITY).
 DR ENBL; AL162754; CAB84224.1; -;
 DR F.R.; A81942; A81942.
 DR GO: 0016491; F: oxidoreductase activity; IEA.
 DR GO: 0006118; P: electron transport; IEA.

DR InterPro: IPR003042; RING_mnxyogenase.
 DR PRINTS: PRO0420; RINGMNOXGNASE.
 KW FAD; Flavoprotein; Hypothetical protein; Oxidoreductase;
 KW Complete proteome.
 SQ SEQUENCE 539 AA; 59518 MW; 3EABCC7E2672BC06 CRC64;

Query Match 66.7%; Score 40; DB 16; Length 539;
 Best Local Similarity 62.5%; Pred. No. 1.7e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 WEYSVWLS 9
 DB 271 WGVSVWLN 278

RESULT 11
 Q86BG4 PRELIMINARY; PRT; 234 AA.

AC Q86BG4
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE CG33169-PB.
 GN CG33169.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfankuch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fessler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Idali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinart K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.

RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochownik S.E., Smith C.D.,
 RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celniker S.E.,
 RA Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
 RA Harris N.L., Krommiller B., Marshall B., Millburn G.H., Richter J.,
 RA Russo S., Searle S.M.J., Smith E., Shu S., Smutniak F.,
 RA Whiffeld E.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,
 RA Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003599; AAC41287.1; -;
 SQ SEQUENCE 234 AA; 26927 MW; 8AF93CE6EF54DF8B CRC64;
 Query Match 65.0%; Score 39; DB 5; Length 234;
 Best Local Similarity 50.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SWEYSVWLSN 10
 |||:|:
 Db 71 SNEFWGQN 80
 RESULT 12
 ID O46331 PRELIMINARY; PRT; 262 AA.
 AC O46331;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Plasmid G1e4293, complete sequence.
 OS Gracilariopsis lemaneiformis.
 OG Plasmid G1e4293.
 OC Eukaryota; Rhodophyta; Florideophyceae; Gracilariaceae; Gracilariopsis;
 OC Gracilariopsis.
 OX NCBI_TaxID=2782;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Goff L.J., Moon D.A.;
 RT "Plasmids of the red algae Gracilaria and Gracilariopsis
 RT (Gracilariaceae): Molecular characterization and cellular
 RT localization.";
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF034720; AAC04749.1; -;
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
 KW Plasmid.
 SQ SEQUENCE 262 AA; 29467 MW; D367CF975975ECFCF CRC64;
 Query Match 65.0%; Score 39; DB 10; Length 262;
 Best Local Similarity 71.4%; Pred. No. 1.2e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 WEYSVWL 8
 ||:||||
 Db 204 WDFSVWL 210
 RESULT 13
 ID Q8DUU7 PRELIMINARY; PRT; 277 AA.
 AC Q8DUU7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Conserved hypothetical protein.
 GN SMU.796.

OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=UA159 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186;
 RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
 RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 RT pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 DR EMBL; AE014920; AAN59514.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 277 AA; 30619 MW; BAE83D5A81904DF CRC64;
 Query Match 65.0%; Score 39; DB 16; Length 277;
 Best Local Similarity 62.5%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 2 WEYSVWLS 9
 |||:|:
 Db 234 WEYDAWS 241
 RESULT 14
 ID Q9J4N2 PRELIMINARY; PRT; 300 AA.
 AC Q9J4N2;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Envelope glycoprotein (Fragment).
 GN ENV.
 OS Human immunodeficiency virus 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=99MX.Y06;
 RX MEDLINE=21094707; PubMed=11177387;
 RA Rivera-Morales L.G., Novitsky V.A., Trujillo J.R.,
 RA Lavallo-Montalvo C., Cano-Dominguez C., Ramos-Jimenez J.,
 RA Jimenez-Rios E., Flores-Flores L., Lopez-Guillen P., Gilbert P.,
 RA Vanberg F., Ramez-Guerra R., Rodriguez-Padilla C., Essex M.;
 RT "The molecular epidemiology of HIV type 1 of men in Mexico.";
 RL AIDS Res. Hum. Retroviruses 17:87-92(2001).
 DR EMBL; AF200914; AAF27424.1; -;
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000777; GPI20.
 DR Pfam; PF00516; GPI20; 1.
 KW AIDS; Coat protein; Glycoprotein.
 FT NON_TER 1
 FT NON_TER 300 300
 SQ SEQUENCE 300 AA; 33175 MW; 0A77AF649C9356C9 CRC64;
 Query Match 65.0%; Score 39; DB 15; Length 300;
 Best Local Similarity 60.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 SWEYSVWLSN 10
 |||:|:
 Db 181 SWNSYNTGQN 190
 RESULT 15
 ID P9S129 PRELIMINARY; PRT; 406 AA.
 AC P9S129; Q50454;

DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein RV2963.
GN RV2963 OR MTCY349.24C OR MT3039.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaiia F.,
RA Badcock K., Basham D., Brown D., Chillingworth I., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R.,
RA Sultston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Smith D.R.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Robison K.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / OshKosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z83018; CAB05414.1; -.
DR EMBL; U00024; AAA50944.1; -.
DR EMBL; AE007125; AAK47365.1; -.
DR PIR; H70670; H70670.
DR TIGR; MT3039; -.
DR TubercuList; RV2963; -.
DR InterPro; IPR005524; DUF318.
DR Pfam; PF03773; DUF318; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 406 AA; 43728 MW; DDA7B4134FFC4D80 CRC64;

Query Match 65.0%; Score 39; DB 16; Length 406;
Best Local Similarity 55.6%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SWEYSVWLS 9
Db 352 SWNTTWIN 360

Search completed: May 19, 2004, 15:46:23
Job time : 22.8333 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 15:29:47 ; Search time 4.79167 Seconds
(without alignments)
108.668 Million cell updates/sec

Title: US-10-005-684-5
Perfect score: 60
Sequence: 1 SWEYSVWLSN 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------|---------------------|
| 1 | 39 | 65.0 | 219 | 1 ERD2 YEAST | P18414 saccharomyc |
| 2 | 38 | 63.3 | 547 | 1 YAGM ECOLI | P77694 escherichia |
| 3 | 37 | 61.7 | 221 | 1 ERD2 PLAF | P33948 plasmodium |
| 4 | 37 | 61.7 | 1849 | 1 IGA4 HAEIN | P45386 haemophilus |
| 5 | 36 | 60.0 | 123 | 1 YUGM BACSU | O05245 bacillus su |
| 6 | 36 | 60.0 | 161 | 1 PADC BACSU | O07006 bacillus su |
| 7 | 36 | 60.0 | 212 | 1 ERD2 SCHPO | O94270 schizosacch |
| 8 | 36 | 60.0 | 216 | 1 T2BB BACSU | P33562 bacillus su |
| 9 | 36 | 60.0 | 752 | 1 PSAA CYACA | O97145 cyanidium c |
| 10 | 36 | 60.0 | 752 | 1 PSAA GUITH | O78508 guillardia |
| 11 | 36 | 60.0 | 826 | 1 OPGH WIGBR | O843d5 wigglewort |
| 12 | 36 | 60.0 | 875 | 1 SYV BORBU | O51680 borrelia bu |
| 13 | 36 | 60.0 | 1195 | 1 CH34 NEUCR | O01285 neurospora |
| 14 | 36 | 60.0 | 1541 | 1 IGA1 HAEIN | P42782 haemophilus |
| 15 | 36 | 60.0 | 1545 | 1 IGA3 HAEIN | P45385 haemophilus |
| 16 | 36 | 60.0 | 1694 | 1 IGA0 HAEIN | P44969 haemophilus |
| 17 | 36 | 60.0 | 1702 | 1 IGA2 HAEIN | P45384 haemophilus |
| 18 | 35.5 | 59.2 | 469 | 1 LET1 KLULA | P53398 kluyveromyc |
| 19 | 35 | 58.3 | 113 | 1 YG40 YEAST | P53307 saccharomyc |
| 20 | 35 | 58.3 | 119 | 1 YB20 YEAST | P38227 saccharomyc |
| 21 | 35 | 58.3 | 190 | 1 NUSM ARBLI | Q33753 arabacia lix |
| 22 | 35 | 58.3 | 219 | 1 ERD2 KLULA | P18413 kluyveromyc |
| 23 | 35 | 58.3 | 413 | 1 YG18 AQUAE | O67545 aquifex aeo |
| 24 | 35 | 58.3 | 791 | 1 VD05 FOWPV | P21969 fowlpox vir |
| 25 | 35 | 58.3 | 839 | 1 CYA1 HUMAN | O08528 homo sapien |
| 26 | 35 | 58.3 | 1064 | 1 CYA4 RAT | P26770 rattus norv |
| 27 | 35 | 58.3 | 1077 | 1 CYA4 HUMAN | Q8nfm4 homo sapien |
| 28 | 35 | 58.3 | 1077 | 1 CYA4 MOUSE | Q91wif3 mus musculu |
| 29 | 35 | 58.3 | 1078 | 1 CYA7 BOVIN | Q29450 bos taurus |
| 30 | 35 | 58.3 | 1080 | 1 CYA7 HUMAN | P51828 homo sapien |
| 31 | 35 | 58.3 | 1090 | 1 CYA2 RAT | P26769 rattus norv |
| 32 | 35 | 58.3 | 1091 | 1 CYA2 HUMAN | Q08462 homo sapien |
| 33 | 35 | 58.3 | 1099 | 1 CYA7 MOUSE | P51829 mus musculu |

| | | | | | |
|----|----|------|------|--------------|--------------------|
| 34 | 35 | 58.3 | 1134 | 1 CVA1 BOVIN | P19754 bos taurus |
| 35 | 35 | 58.3 | 1144 | 1 CVA3 HUMAN | O60266 homo sapien |
| 36 | 35 | 58.3 | 1144 | 1 CVA3 RAT | P21932 rattus norv |
| 37 | 35 | 58.3 | 1145 | 1 CVA3 MOUSE | Q8vnh7 mus musculu |
| 38 | 34 | 56.7 | 146 | 1 YDGK ECOLI | P76180 escherichia |
| 39 | 34 | 56.7 | 227 | 1 SRPB SCHPO | O13950 schizosacch |
| 40 | 34 | 56.7 | 253 | 1 YDIJ BACHD | Q929p4 bacillus ha |
| 41 | 34 | 56.7 | 294 | 1 PFTA SCHPO | O60052 s protein f |
| 42 | 34 | 56.7 | 315 | 1 OWPP ECOLI | P34210 escherichia |
| 43 | 34 | 56.7 | 326 | 1 T2PS PROST | F00640 providencia |
| 44 | 34 | 56.7 | 351 | 1 Y887 METTH | O26973 methanobact |
| 45 | 34 | 56.7 | 384 | 1 CARA PSEST | P38099 pseudomonas |

ALIGNMENTS

RESULT 1
ERD2 YEAST
ID ERD2 YEAST STANDARD; PRT; 219 AA.
AC P18414;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ER lumen protein retaining receptor (HDEL receptor).
GN ERD2 OR YBL040C OR YBL0408.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEY2102;
RX MEDLINE=90304893; PubMed=2194670;
RA Semenza J.C., Hardwick K.G., Dean N., Pelham H.R.B.;
RT "ERD2, a yeast gene required for the receptor-mediated retrieval of
luminal ER proteins from the secretory pathway."
RL Cell 61.1349-1357(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=95176707; PubMed=7871888;
RA de Wergifosse P., Jacques B., Jonniaux J.-L., Purnelle B., Skala J.,
Goffeau A.;
RT "The sequence of a 22.4 kb DNA fragment from the left arm of yeast
chromosome II reveals homologues to bacterial proline synthetase and
murine alpha-adaptin, as well as a new permease and a DNA-binding
protein."
RL Yeast 10.1489-1496(1994).
CC -!- FUNCTION: Required for the retention of luminal endoplasmic
reticulum proteins. Determines the specificity of the luminal ER
protein retention system. Also required for normal vesicular
traffic through the Golgi. This receptor strongly recognizes
H-D-E-L and weakly recognizes D-D-E-L and K-D-E-L.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the ERD2 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M34777; AAA68907.1; -;
CC EMBL; X78214; CAA55054.1; ALT_INIT.
CC EMBL; Z35801; CAA84860.1; -;
CC PIR; A35617; A35617;
CC GeneOnline; 138475; -;
CC SGD; S0000136; ERD2.
CC GO; GO:0030176; C:integral to endoplasmic reticulum membrane; IDA.
CC GO; GO:0045015; F:HDEL sequence binding; IGI.

DR GO: 0006621; P: protein-ER retention; IGI.
DR InterPro: IPR000133; ERret_receptor.
DR Pfam: PF00810; ER_lumen_recept; 1.
DR PRINTS: PR00660; ERLUMENR.
DR ProDom: PD005774; ERret_receptor; 1.
DR PROSITE: PS00951; ER_LUMEN RECEPTOR 1; 1.
DR PROSITE: PS00952; ER_LUMEN RECEPTOR 2; 1.
KW Endoplasmic reticulum; Transmembrane; Protein transport; Receptor.
FT DOMAIN 1 2 LUMENAL (POTENTIAL).
FT TRANSMEM 3 21 POTENTIAL.
FT DOMAIN 22 35 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 36 53 POTENTIAL.
FT DOMAIN 54 62 LUMENAL (POTENTIAL).
FT TRANSMEM 63 82 POTENTIAL.
FT DOMAIN 83 102 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 103 116 POTENTIAL.
FT DOMAIN 117 123 LUMENAL (POTENTIAL).
FT TRANSMEM 124 143 POTENTIAL.
FT DOMAIN 144 155 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 156 174 POTENTIAL.
FT DOMAIN 175 185 LUMENAL (POTENTIAL).
FT TRANSMEM 186 206 POTENTIAL.
FT DOMAIN 207 219 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 219 AA; 25762 MW; 9F19C47F81C40182 CRC64;

Query Match 65.0%; Score 39; DB 1; Length 219;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SWEYSVWLSN 10
Db 125 AWSFSVWLS 134
:|:||||:
125 AWSFSVWLS 134

RESULT 2
YAGW ECOLI
ID YAGW_ECOLI STANDARD; PRT; 547 AA.
AC P77694;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yagW.
GN YAGW OR B0290.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KL2 / MGI655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Durcan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
RA Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RA Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF000136; AAC73393.1; -.
DR

DR EMBL; U73857; AAB18019.1; -.
DR PIR; B64755; B64755.
DR RCGene; EG13562; yagW.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 547 AA; 59997 MW; DF6467ECD8785E3 CRC64;

Query Match 63.3%; Score 38; DB 1; Length 547;
Best Local Similarity 40.0%; Pred. No. 60;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 SWEYSVWLSN 10
Db 97 NWKFDWLEN 106
:|:|:|:
97 NWKFDWLEN 106

RESULT 3
ERD2_PLAFA
ID ERD2_PLAFA STANDARD; PRT; 221 AA.
AC P33948;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE ER lumen protein retaining receptor.
GN ERD2.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ITGZPF6;
RX MEDLINE=94038960; PubMed=8223485;
RA Elmendorf H.G., Haldar K.;
RT Identification and localization of ERD2 in the malaria parasite
RT Plasmodium falciparum: separation from sites of sphingomyelin
RT synthesis and implications for organization of the Golgi.";
RL EMBO J. 12:4763-4773 (1993).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the ERD2 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X74869; CAAS2861.1; -.
DR EMBL; Z26043; CAAS1128.1; -.
DR PIR; S39609; S39609.
DR InterPro: IPR000133; ERret_receptor.
DR Pfam: PF00810; ER_lumen_recept; 1.
DR PRINTS: PR00660; ERLUMENR.
DR ProDom: PD005774; ERret_receptor; 1.
DR PROSITE: PS00951; ER_LUMEN RECEPTOR 1; 1.
DR PROSITE: PS00952; ER_LUMEN RECEPTOR 2; 1.
KW Endoplasmic reticulum; Transmembrane; protein transport; Receptor.
FT DOMAIN 1 2 LUMENAL (POTENTIAL).
FT TRANSMEM 3 21 POTENTIAL.
FT DOMAIN 22 35 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 36 52 POTENTIAL.
FT DOMAIN 53 61 LUMENAL (POTENTIAL).
FT TRANSMEM 62 80 POTENTIAL.
FT DOMAIN 81 99 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 100 113 POTENTIAL.
FT DOMAIN 114 120 LUMENAL (POTENTIAL).
FT TRANSMEM 121 140 POTENTIAL.
FT DOMAIN 141 152 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 153 171 POTENTIAL.
FT DOMAIN 172 183 LUMENAL (POTENTIAL).
FT TRANSMEM 184 203 POTENTIAL.
FT DOMAIN 204 221 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 221 AA; 26446 MW; 7F231ED6AAACB368 CRC64;

Query Match 61.7%; Score 37; DB 1; Length 221;
Best Local Similarity 44.4%; Pred. No. 34;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 WEYSVWLSN 10
| :|:| :
DB 123 WFSFVWLSN 131

RESULT 4
IGA4 HAEIN
ID IGA4 HAEIN STANDARD; PRT; 1849 AA.
AC P45386;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Immunoglobulin A1 protease precursor (EC 3.4.21.72) (IGA1 protease).
GN IGA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN SEQUENCE FROM N.A.
RP STRAIN=NH71 HK61;
RX MEDLINE=92234949; PubMed=1373717;
RA Poulson K., Reinholdt J., Killian M.;
RT "A comparative genetic study of serologically distinct Haemophilus
influenzae type 1 immunoglobulin A1 proteases.";
RL J. Bacteriol. 174:2913-2921(1992).
CC -!- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
CC -!- PRODUCING INTRAF. FC AND FAB FRAGMENTS.
CC -!- CATALYTIC ACTIVITY: Cleavage of immunoglobulin A molecules at
CC certain Pro-Xaa bonds in the hinge region. No small molecule
CC substrates are known.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
CC SIMILARITY).
CC -!- SIMILARITY: Belongs to peptidase family S6.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M87491; AAA24968.1; -;
CC MEROPS; S06.001; -;
CC InterPro; IPR006315; Autotransport.
CC InterPro; IPR005546; Autotransporter.
CC InterPro; IPR000710; Peptidase_S6.
CC InterPro; IPR004899; Pertactin.
CC Pfam; PF03797; Autotransporter; 1.
CC Pfam; PF02395; IGA1; 1.
CC Pfam; PF03212; Pertactin; 1.
CC PRINTS; PR00921; IGASERPTASE.
CC TIGRfams; TIGR01414; autotrans_bar1; 1.
KW Hydrolase; Serine protease; Transmembrane; Zymogen; Signal.
FT SIGNAL
FT CHAIN 1 25 POTENTIAL.
FT PROPEP 1022 1849 IMMUNOGLOBULIN A1 PROTEASE.
FT ACT_SITE 299 299 HELPER PEPTIDE (POTENTIAL).
FT SEQUENCE 1849 AA; 202957 MW; 79A7D018C7150AEA CRC64;
SQ

Query Match 61.7%; Score 37; DB 1; Length 1849;
Best Local Similarity 62.5%; Pred. No. 3e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 EYSVWLSN 10
| :|:| :
DB 1601 QYNVWLSN 1608

RESULT 5
YUGM BACSU
ID YUGM BACSU STANDARD; PRT; 123 AA.
AC O05245;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein yugM.
GN YUGM OR BSU31340.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN SEQUENCE FROM N.A.
RP STRAIN=168;
RX MEDLINE=97419515; PubMed=9274030;
RA Oudega B., Konings G., Rodrigues L., de Sales Ramon M.,
RA Hilbert H., Puesterhoef A., Pohl T.M., Weitzenecker T.;
RT "Analysis of the Bacillus subtilis genome: cloning and nucleotide
RT sequence of a 62 kb region between 275 degrees (rnbB) and 284 degrees
RT (pai)";
RL Microbiology 143:2769-2774 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Bron S.,
RA Broussier R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Bouilliet S., Bruchi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Ertan K.D., Ertington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghm S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Konings G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F., Soldo B.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Taccini E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenberg M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis";
RL Nature 390:249-256 (1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```

DR  ENBL; Z93936; CAB07931.1; -.
DR  ENBL; Z99120; CAB15123.1; -.
DR  PIR; C70011; C70011.
DR  Subtilisin; BG12367; yugM.
KW  Hypothetical protein; Transmembrane; Complete proteome.
FT  TRANSMEM 7 29 Potential.
FT  TRANSMEM 44 66 Potential.
FT  TRANSMEM 79 101 Potential.
SQ  SEQUENCE 123 AA; 13686 MW; CDAZADBOE493EA52 CRC64;

Query Match 60.0%; Score 36; DB 1; Length 123;
Best Local Similarity 55.6%; Pred. No. 27;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWEYSWLS 9
DB 89 SWDSAVWLA 97

RESULT 6
PADC_BACSU STANDARD; PRT; 161 AA.
AC 007006;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phenolic acid decarboxylase (EC 4.1.1.-) (PAD).
GN PADC OR PAD OR BSU34400.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=168;
RX MEDLINE=98207851; PubMed=9546183;
RA Cavin J.-F., Dartois V., Davies C.;
RT "Gene cloning, transcriptional analysis, purification, and
RT characterization of phenolic acid decarboxylase from Bacillus
RT subtilis.";
RL Appl. Environ. Microbiol. 64:1466-1471(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,

```

```

RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vasearotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC -!- FUNCTION: Catalyzes the decarboxylation of phenolic acids such as
CC ferulic, p-coumaric and caffeic acids.
CC -!- SUBUNIT: Homodimer (Probable).
CC -!- INDUCTION: By ferulic, p-coumaric and caffeic acids. Cells
CC extracts from caffeic acid-induced cells exhibited lower activity
CC on the three acids, which indicates that caffeic acid could be a
CC less efficient inducer.
CC -!- MISCELLANEOUS: Optimal temperature is 40-45 degrees Celsius and
CC optimal pH is 5.0.
CC -!- SIMILARITY: BELONGS TO THE PADC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  ENBL; AF017117; AAC46254.1; -.
DR  ENBL; Z94043; CAB08020.1; -.
DR  ENBL; Z99121; CAB15445.1; -.
DR  PIR; D69671; D69671.
DR  Subtilisin; BG12433; padC.
DR  InterPro; IPR008729; PA decarbox.
DR  Pfam; PF05870; PA decarbox; 1.
KW  Lyase; Decarboxylase; Complete proteome.
SQ  SEQUENCE 161 AA; 19076 MW; BAF73F679D0FC313 CRC64;

Query Match 60.0%; Score 36; DB 1; Length 161;
Best Local Similarity 44.4%; Pred. No. 36;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 WEYSWLSN 10
DB 17 WEYIYIKN 25

RESULT 7
ERD2_SCHPO STANDARD; PRT; 212 AA.
AC 094270;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ER lumen protein retaining receptor.
GN SPBP8B7.22.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859350;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Horsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

```


RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Wiltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzys K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: Required for the retention of luminal endoplasmic
CC reticulum proteins. Determines the specificity of the luminal ER
CC protein retention system. Also required for normal vesicular
CC traffic through the Golgi (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the ERD2 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL032684; CRA21807.1; -;
CC PIR; T40816; T40816;
CC GeneDB_SF006; SPB8B7.22; -;
CC InterPro; IPR000133; Errret_receptor.
CC Pfam; PF00810; ER_lumen_recept; 1.
CC PRINTS; PR00660; ERLUMENR.
CC ProDom; PD005774; Errret_receptor; 1.
CC PROSITE; PS00951; ER_LUMEN_RECEPTOR_1; 1.
CC PROSITE; PS00952; ER_LUMEN_RECEPTOR_2; 1.
CC
CC Endoplasmic reticulum; Transmembrane; Protein transport; Receptor.
FT DOMAIN 1 2
FT LUMENAL (POTENTIAL).
FT TRANSMEM 3 21
FT POTENTIAL.
FT DOMAIN 22 35
FT CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 36 53
FT POTENTIAL.
FT DOMAIN 54 61
FT LUMENAL (POTENTIAL).
FT TRANSMEM 62 80
FT POTENTIAL.
FT DOMAIN 81 96
FT CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 97 110
FT POTENTIAL.
FT DOMAIN 111 117
FT LUMENAL (POTENTIAL).
FT TRANSMEM 118 137
FT POTENTIAL.
FT DOMAIN 138 149
FT CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 150 168
FT POTENTIAL.
FT DOMAIN 169 178
FT LUMENAL (POTENTIAL).
FT TRANSMEM 179 199
FT POTENTIAL.
FT DOMAIN 200 212
FT CYTOPLASMIC (POTENTIAL).
FT SEQUENCE 212 AA; 24724 MW; 633F1E598D5DC83D CRC64;
Query Match 60.0%; Score 36; DB 1; Length 212;
Best Local Similarity 44.4%; Pred. No. 48;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 2 WEYSVWLSN 10
DB 120 WTSFVWLSN 128
RESULT 8
T2BB_BACSU STANDARD; PRT; 316 AA.
ID T2BB_BACSU
AC P33562;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
CC

DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Type II restriction enzyme BsuBI (EC 3.1.21.4) (Endonuclease BsuBI)
DE (R.BauBI)
GN HSDR OR HSRB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISB8;
RX MEDLINE=93126092; PubMed=1480472;
RA Xu G.-L., Kapfer W., Walter J., Trautner T.A.;
RT "BsuBI -- an isospecific restriction and modification system of PstI:
RT characterization of the BsuBI genes and enzymes.";
RL Nucleic Acids Res. 20:6517-6523(1992).
CC -!- FUNCTION: RECOGNIZES THE DOUBLE-STRANDED SEQUENCE CTGCAG AND
CC CLEAVES AFTER A-S.
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage of DNA to give
CC specific double-stranded fragments with terminal 5'-phosphates.
CC -!- COFACTOR: Magnesium.
CC -!- SUBUNIT: Homodimer.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L01541; AAA18170.1; -;
CC PIR; S35516; S35516.
CC REBASE; 616; BsuBI.
CC Restriction system; Hydrolase; Nuclease; Endonuclease; Magnesium.
CC SEQUENCE 316 AA; 36197 MW; B1C2E45851E7F3C9 CRC64;
Query Match 60.0%; Score 36; DB 1; Length 316;
Best Local Similarity 50.0%; Pred. No. 72;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 SWEYSVWLSN 10
DB 289 SWTEVWVAD 298
RESULT 9
PSAA_CYACA STANDARD; PRT; 752 AA.
ID PSAA_CYACA
AC Q9TLQ5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Photosystem I P700 chlorophyll A apoprotein A1 (PsaA) (PSI-A).
GN PSAA.
OS Cyanidium caldarium.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
OC Cyanidium.
OX NCBI_TaxID=2771;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RX-1;
RX MEDLINE=20496959; PubMed=11040290;
RA Gloeckner G., Rosenthal A., Valentin K.-U.;
RT "The structure and gene repertoire of an ancient red algal plastid
RT genome.";
RL J. Mol. Evol. 51:382-390(2000).
CC -!- FUNCTION: PsaA and psaB bind P700, the primary electron donor of
CC photosystem I (PSI), as well as the electron acceptors A0, A1 and
CC FX. PSI is a plastocyanin/cytochrome c6-ferredoxin oxidoreductase,
CC converting photonic excitation into a charge separation, which
CC transfers an electron from the donor P700 chlorophyll pair to the
CC spectroscopically characterized acceptors A0, A1, FX, FA and FB in


```

Best Local Similarity 50.0%; Pred.No. 1.9e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 SWEYSVWLSN 10
Db      173 SWSYTIYLFN 182

RESULT 12
SYV_BORBU STANDARD; PRT; 875 AA.
ID _SYV_BORBU
AC C51580;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE VALS-tRNA synthetase (EC 6.1.1.9) (Valine-tRNA ligase) (ValRS).
GN VALS OR BB0738.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria, Spirochaetes, Spirochaetales, Spirochaetaceae, Borrelia.
OX NCBI_TaxID=139;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Karpavich A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RL Nature 390:580-586 (1997).
CC -I- CATALYTIC ACTIVITY: ATP + L-valine + tRNA(Val) = AMP + diphosphate
CC + L-valyl-tRNA(Val).
CC -I- SUBUNIT: Monomer (By similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC -I- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001173; AAC67078.1; -.
CC PIR; A70192; A70192.
CC HSP; P96142; LGAX.
CC TIGR; BB0738; -.
DR InterPro; IPR002300; tRNA-synt_1a.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002303; tRNA-synt_1l.
DR InterPro; IPR009008; ValRS lIeRS_edit.
DR Pfam; PF001133; tRNA-synt_1; 1.
DR PRINTS; PR00986; TRNASYNTHVAL.
DR TIGRFAMs; TIGR00422; vals; 1.
DR PROSITE; PS00178; AA tRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 45 55 "HIGH" REGION.
FT SITE 524 528 "KMSKS" REGION.
FT BINDING 527 527 ATP (BY SIMILARITY).
SQ SEQUENCE 875 AA; 101864 MW; D9EF63611A4BC77E CRC64;

Query Match 60.0%; Score 36; DB 1; Length 875;
Best Local Similarity 63.6%; Pred.No. 2e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 2; Gaps 1

QY      2 WE--YSVWLSN 10

```

```

Db          || | | | |
            387 WENTYKYLNSN 397

RESULT 13
CHS4_NEUCR
ID CHS4_NEUCR STANDARD; PRT; 1195 AA.
AC Q01285;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chitin synthase 4 (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl
DE transferase 4) (Class-IV chitin synthase 4).
GN CHS-4.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RX MEDLINE=96188842; PubMed=8628221;
RA Din A.B., Specht C.A., Robbins F.W., Yarden O.;
RT "chs-4, a class IV chitin synthase gene from Neurospora crassa.";
RL Mol. Gen. Genet. 250:214-222(1996).
CC -1- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + {(1,4)-(N-acetyl-
CC beta-D-glucosaminyl)}(N) = UDP + {(1,4)-(N-acetyl-beta-D-
CC glucosaminyl)}(N+1).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane.
CC -1- SIMILARITY: Belongs to the chitin synthase family. Subfamily class
CC IV.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U25097; AAB03563.1; -.
CC PIR; S61886; S61886.
CC InterPro; IPR004835; Chitin_synth_fng.
CC InterPro; IPR001199; Cyt_B5.
CC Pfam; PF031142; Chitin_synth_2; 1.
CC Pfam; PF00173; heme_1; 1.
CC Transferrase; Glycosyltransferase; Transmembrane; Cell wall;
CC Multigene family.
CC TRANSMEM 205 225 POTENTIAL.
CC TRANSMEM 475 495 POTENTIAL.
CC TRANSMEM 885 905 POTENTIAL.
CC TRANSMEM 1026 1046 POTENTIAL.
CC TRANSMEM 1053 1073 POTENTIAL.
CC TRANSMEM 1077 1097 POTENTIAL.
CC SEQUENCE 1195 AA; 133544 MW; 353F211FABEF1D30 CRC64;

Query Match 60.0%; Score 36; DB 1; Length 1195;
Best Local Similarity 50.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SWEYSVWL 8
    ||| :|
    1076 SWSYIWM 1083

Db          || | | | |
            387 WENTYKYLNSN 397

RESULT 14
IGAL_HAEIN
ID IGAL_HAEIN STANDARD; PRT; 1541 AA.
AC P42782;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Immunoglobulin A1 protease precursor (EC 3.4.21.72) (IGAL protease).

```

```

GN IGA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurelliales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HK368 / Serotype B;
RX MEDLINE=89379374; PubMed=2506130;
RA Poulsen K., Brandt J., Hjorth J.P., Thoegersen H.C., Kilian M.;
RT "Cloning and sequencing of the immunoglobulin A1 protease gene (iga)
RT of Haemophilus influenzae serotype b.";
RL Infect. Immun. 57:3097-3105(1989).
RN [2]
RP MUTAGENESIS OF SER-288.
RC STRAIN=HK368 / Serotype B;
RX MEDLINE=92234949; PubMed=1373717;
RA Poulsen K., Reinholdt J., Kilian M.;
RT "A comparative genetic study of serologically distinct Haemophilus
RT influenzae type 1 immunoglobulin A1 proteases.";
RL J. Bacteriol. 174:2913-2921(1992).
CC -1- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
CC -1- PRODUCING INTACT FC AND FAB FRAGMENTS.
CC -1- CATALYTIC ACTIVITY: Cleavage of immunoglobulin A molecules at
CC certain Pro-|-Xaa bonds in the hinge region. No small molecule
CC substrates are known.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
CC SIMILARITY).
CC -1- SIMILARITY: Belongs to peptidase family S6.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X64357; CAA45708.1; -.
CC EMBL; M87492; AAA24969.1; -.
CC PIR; A37023; A37023.
CC MEROPS; S06.001; -.
CC InterPro; IPR006315; Autotransport.
CC InterPro; IPR005546; Autotransporter.
CC InterPro; IPR000710; Peptidase_S6.
CC InterPro; IPR004899; Pertactin.
CC Pfam; PF03797; Autotransporter; 1.
CC Pfam; PF02395; IGA1; 1.
CC Pfam; PF03212; Pertactin; 1.
CC PRINTS; PR00921; IGASERPTASE.
CC TIGRFS; TIGR01414; autotrans_bar1; 1.
KW Hydrolyase; Serine protease; Transmembrane; Zymogen; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 1008 IMMUNOGLOBULIN A1 PROTEASE.
FT PROPEP 1009 1541 HELPER PEPTIDE (POTENTIAL).
FT ACT_SITE 288 288 PROBABLE.
FT MUTAGEN 288 288 S->T; LOSS OF ACTIVITY.
SQ SEQUENCE 1541 AA; 169370 MW; CE7257CH3196C600 CRC64;

Query Match 60.0%; Score 36; DB 1; Length 1541;
Best Local Similarity 62.5%; Pred. No. 3.7e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 EYSVWLSN 10
    :|:|:|
    1293 QYNVWVSN 1300

Db          || | | | |
            387 WENTYKYLNSN 397

RESULT 15

```

IGA3 HAEIN
ID IGA3 HAEIN STANDARD; PRT; 1545 AA.
AC P45385;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Immunoglobulin A1 protease precursor (EC 3.4.21.72) (IGA1 protease).
GN IGA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HK393 / NCTC 8467 / Serotype B;
RX MEDLINE=92234949; PubMed=1373717;
RA Poulson K., Reinholdt J., Killian M.;
RT "A comparative genetic study of serologically distinct Haemophilus
RT influenzae type 1 immunoglobulin A1 proteases.";
RL J. Bacteriol. 174:2913-2921(1992).
CC -!- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
CC -!- PRODUCING INTACT FC AND FAB FRAGMENTS.
CC -!- CATALYTIC ACTIVITY: Cleavage of immunoglobulin A molecules at
CC certain Pro-|-Xaa bonds in the hinge region. No small molecule
CC substrates are known.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
CC SIMILARITY).
CC -!- SIMILARITY: Belongs to peptidase family S6.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M87490; AAA24967.1; -.
DR FTR; B41859; B41859.
DR MEROPS; S06.001; -.
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR005546; Autotransporter.
DR InterPro; IPR000710; Peptidase_S6.
DR InterPro; IPR004899; Pertactin.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF02395; IGA1; 1.
DR Pfam; PF03212; Pertactin; 1.
DR PRINTS; PR00921; IGASERPTASE.
DR TIGRFAMs; TIGR01414; autotrans_barl; 1.
KW Hydrolyase; Serine protease; Transmembrane; Zymogen; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 1012 IMMUNOGLOBULIN A1 PROTEASE.
FT PROPEP 1013 1545 HELPER PEPTIDE (POTENTIAL).
FT ACT SITE 292 292 PROBABLE.
SQ SEQUENCE 1545 AA; 170627 MW; 3EDD753988F6D478 CRC64;

Query Match 60.0%; Score 36; DB 1; Length 1545;
Best Local Similarity 62.5%; Pred. No. 3.7e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 EYSVWLSN 10
DB 1297 QYNWWSN 1304

Search completed: May 19, 2004, 15:44:29
Job time : 5.79167 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 15:39:22 ; Search time 7.29167 Seconds
(without alignments)
131.920 Million cell updates/sec

Title: US-10-005-684-5

Perfect score: 60

Sequence: 1 SWEYSWLSN 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 41 | 68.3 | 319 | 2 T27118 | hypothetical prote |
| 2 | 40 | 66.7 | 539 | 2 A81942 | hypothetical prote |
| 3 | 39 | 65.0 | 219 | 1 A35617 | HDEL receptor ERD2 |
| 4 | 39 | 65.0 | 406 | 2 H70670 | hypothetical prote |
| 5 | 38 | 63.3 | 77 | 2 G90429 | hypothetical prote |
| 6 | 38 | 63.3 | 135 | 2 D90348 | hypothetical prote |
| 7 | 38 | 63.3 | 148 | 2 A99348 | hypothetical prote |
| 8 | 38 | 63.3 | 185 | 2 AB2241 | hypothetical prote |
| 9 | 38 | 63.3 | 212 | 2 AE1473 | hypothetical prote |
| 10 | 38 | 63.3 | 251 | 2 A11111 | hypothetical prote |
| 11 | 38 | 63.3 | 288 | 2 F90331 | partial transposas |
| 12 | 38 | 63.3 | 310 | 2 E90334 | partial transposas |
| 13 | 38 | 63.3 | 332 | 2 C90497 | transposase ISC135 |
| 14 | 38 | 63.3 | 350 | 2 A90279 | transposase ISC135 |
| 15 | 38 | 63.3 | 363 | 2 C90197 | transposase ISC135 |
| 16 | 38 | 63.3 | 375 | 2 B90339 | transposase ISC135 |
| 17 | 38 | 63.3 | 388 | 2 H90302 | transposase ISC135 |
| 18 | 38 | 63.3 | 547 | 2 B64755 | yagW protein - Esc |
| 19 | 38 | 63.3 | 547 | 2 H90668 | probable receptor |
| 20 | 38 | 63.3 | 547 | 2 C85519 | potassium uptake p |
| 21 | 38 | 63.3 | 639 | 2 A12663 | potassium uptake p |
| 22 | 38 | 63.3 | 639 | 2 H97445 | beta 1,3 glucan sy |
| 23 | 38 | 63.3 | 654 | 2 AB2932 | hypothetical prote |
| 24 | 38 | 63.3 | 654 | 2 D98350 | probable TonB-depe |
| 25 | 38 | 63.3 | 789 | 2 H83354 | TonB-dependent rec |
| 26 | 38 | 63.3 | 1105 | 2 C87622 | duplicated hypothe |
| 27 | 38 | 63.3 | 772 | 2 C81887 | transcription regu |
| 28 | 37 | 61.7 | 208 | 2 AD2368 | |
| 29 | 37 | 61.7 | 208 | 2 AD2368 | |

30 37 61.7 221 2 S39609 ERD2 protein - mal
31 37 61.7 1849 2 C41859 IGA-specific metal
32 36 60.0 72 2 T42329 hypothetical prote
33 36 60.0 123 2 C70011 hypothetical prote
34 36 60.0 152 2 AG2513 hypothetical prote
35 36 60.0 161 2 D69671 ferulate decarboxy
36 36 60.0 212 2 T40816 ER lumen protein r
37 36 60.0 316 2 S35516 type II site-speci
38 36 60.0 333 2 F70341 sulfur oxygenase x
39 36 60.0 380 2 T48491 giberellin 20-oxi
40 36 60.0 543 2 T01100 reverse transcript
41 36 60.0 657 2 T01301 RNA-directed DNA p
42 36 60.0 716 2 T39812 hypothetical prote
43 36 60.0 875 2 A70192 valine-tRNA ligase
44 36 60.0 1195 2 S61886 chitin synthase (E
45 36 60.0 1541 2 A37023 IGA-specific metal

ALIGNMENTS

RESULT 1

T27118

hypothetical protein Y53C10A.5 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T27118

R;White, S.

submitted to the EMBL Data Library, November 1998

A;Reference number: Z20314

A;Accession: T27118

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-319 <MIL>

A;Cross-references: EMBL:AL033536; PIDN:CAA22139.1; CESP:Y53C10A.5

A;Experimental source: clone Y53C10A

C;Genetics:

A;Gene: CESP:Y53C10A.5

A;Introns: 31/2; 91/1; 124/1; 233/3; 264/3; 290/1

Query Match

Best Local Similarity 68.3%; Score 41; DB 2; Length 319;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWEYSWLS 9

||| |.|

Db 117 SWEYSWLS 125

RESULT 2

A81942

hypothetical protein NMA0954 [imported] - Neisseria meningitidis (strain Z2491 serogroup

C;Species: Neisseria meningitidis

C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C;Accession: A81942

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel;

; Holroyd, S.; Jagels, K.; Leather, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A;Reference number: A81775; MUID:20222556; PMID:10761919

A;Accession: A81942

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-539 <PAR>

A;Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CA84224.1; PID:g737965

A;Experimental source: serogroup A, strain Z2491

C;Genetics:

A;Gene: NMA0954

Query Match

Best Local Similarity 66.7%; Score 40; DB 2; Length 539;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 WEYSVWLS 9

Db 271 WQYGVWLN 278

RESULT 3

HDEL receptor ERD2 - yeast (*Saccharomyces cerevisiae*)

N;Alternate names: 26k endoplasmic reticulum retention receptor; ER lumen protein-retention

C;Species: *Saccharomyces cerevisiae*

C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 21-Jul-2000

C;Accession: A35617; S45774; S50290; S42504

R;Semenza, J.C.; Hardwick, K.G.; Dean, N.; Pelham, H.R.B.

Cell 61, 1349-1357, 1990

A;Title: ERD2, a yeast gene required for the receptor-mediated retrieval of luminal ER P

A;Reference number: A35617; MUID:90304893; PMID:2194670

A;Accession: A35617

A;Molecule type: DNA

A;Residues: 1-219 <SEM>

A;Cross-references: GB:M34777; NID:G171466; PIDN:AAA68907.1; PID:G171467

R;Goffeau, A.; Jonniaux, J.L.; Purnelle, B.; Skala, J.; de Wergifosse, P.; van Dycck, L.

submitted to the Protein Sequence Database, August 1994

A;Reference number: S45774

A;Accession: S45774

A;Molecule type: DNA

A;Residues: 1-219 <GOF>

A;Cross-references: EMBL:Z35801; NID:G536054; PIDN:CAA84860.1; PID:G536055; GSPDB:GN0000

R;de Wergifosse, P.; Jacques, B.; Jonniaux, J.L.; Purnelle, B.; Skala, J.; Goffeau, A.

Yeast 10, 1489-1496, 1994

A;Title: The sequence of a 22.4 kb DNA fragment from the left arm of yeast chromosome II

NA-binding protein.

A;Reference number: S50284; MUID:95176707; PMID:7871888

A;Accession: S50290

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 64-219 <DEF>

A;Cross-references: EMBL:X78214; NID:G463261; PIDN:CAA55054.1; PID:G463268

A;Note: The nucleotide sequence was submitted to the EMBL Data Library, March 1994

C;Genetics:

A;Gene: SGD:ERD2; MIPS:YBL040C

A;Cross-references: SGD:S0000136; MIPS:YBL040C

A;Map position: 2L

A;Introns: 8/1

C;Function:

A;Description: achieves retention of proteins specific to the lumen of the endoplasmic

usually KDEL in animal cells and HDEL in budding yeasts

C;Superfamily: KDEL receptor

F;4-20/Domain: transmembrane #status predicted <TM1>

F;38-54/Domain: transmembrane #status predicted <TM2>

F;61-82/Domain: transmembrane #status predicted <TM3>

F;99-115/Domain: transmembrane #status predicted <TM4>

F;122-138/Domain: transmembrane #status predicted <TM5>

F;186-202/Domain: transmembrane #status predicted <TM6>

QY 1 SWEYSVWLSN 10

Db 125 AWSFSVWLS 134

RESULT 4

Hypothetical protein Rv2963 - *Mycobacterium tuberculosis* (strain H37RV)

N;Alternate names: u0002f protein

C;Species: *Mycobacterium tuberculosis*

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C;Accession: H70670; S73059

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: H70670

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-406 <COL>

A;Cross-references: GB:Z83018; GB:AL123456; NID:G3261671; PIDN:CA805414.1; PID:e283377; I

A;Experimental source: strain H37RV

R;Smith, D.R.; Robinson, K.

submitted to the EMBL Data Library, September 1994

A;Description: *Mycobacterium tuberculosis* cosmid tbcc2.

A;Reference number: S73053

A;Accession: S73059

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-406 <SMI>

A;Cross-references: EMBL:U00024; NID:G560506; PIDN:AAA50944.1; PID:G560523

C;Genetics:

A;Gene: Rv2963

A;Start codon: GTG

Query Match 65.0%; Score 39; DB 2; Length 406;

Best Local Similarity 55.6%; Pred. No. 47;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SWEYSVWLS 9

Db 352 SWNYTTWLN 360

RESULT 5

G90429

hypothetical protein SSO10674 [imported] - *Sulfolobus solfataricus* transposon ISCL359

C;Species: *Sulfolobus solfataricus*

C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C;Accession: G90429

R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-v

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.

submitted to GenBank, April 2001

A;Description: *Sulfolobus solfataricus* complete genome.

A;Reference number: A99139

A;Accession: G90429

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-77 <KUR>

A;Cross-references: GB:AE006641; NID:G13815870; PIDN:AAK42694.1; GSPDB:GN00155

C;Genetics:

A;Gene: SSO10674

Query Match 63.3%; Score 38; DB 2; Length 77;

Best Local Similarity 77.8%; Pred. No. 12;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 WEYSVWLSN 10

Db 30 WEFSVLSN 38

RESULT 6

D90348

hypothetical protein SSO1853 [imported] - *Sulfolobus solfataricus* transposon ISCL359

C;Species: *Sulfolobus solfataricus*

C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C;Accession: D90348

R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-v

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.

submitted to GenBank, April 2001

A;Description: *Sulfolobus solfataricus* complete genome.

A:Reference number: A99139
A:Accession: D90348
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-135 <KUR>
A:Cross-references: GB:AE006641; NID:gl13815107; PIDN:AAK42043.1; GSPDB:GN00155
C:Genetics:
A:Gene: SSO1853

Query Match 63.3%; Score 38; DB 2; Length 135;
Best Local Similarity 77.8%; Pred. No. 21;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 WEYSVWLSN 10
Db 105 WEFSVLSN 113
|||:|||||

RESULT 7

A99348
hypothetical protein SSO1847 [imported] - Sulfolobus solfataricus transposon ISCl359
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: A99348
R:Singh, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.; Jone, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: A99348
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148 <KUR>
A:Cross-references: GB:AE006641; NID:gl13815104; PIDN:AAK42040.1; GSPDB:GN00155
C:Genetics:
A:Gene: SSO1847

Query Match 63.3%; Score 38; DB 2; Length 148;
Best Local Similarity 77.8%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 WEYSVWLSN 10
Db 105 WEFSVLSN 113
|||:|||||

RESULT 8

AB2241
hypothetical protein alr3481 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AB2241
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AB2241
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-185 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA075180.1; PID:gl17132614; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr3481

Query Match 63.3%; Score 38; DB 2; Length 185;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SWEYSVWLS 9

Db 149 SMCYLWLS 157
|||:|||||

RESULT 9

AE1473
hypothetical protein lin0324 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AE1473
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AE1473
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-212 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC95557.1; PID:gl16412753; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0324

Query Match 63.3%; Score 38; DB 2; Length 212;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SWEYSVWLSN 10
Db 113 SWTFSTWVHN 122
|||:|||||

RESULT 10

AE1111
hypothetical protein lmo0296 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AE1111
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AE1111
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-212 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00823.1; PID:gl16409661; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo0296

Query Match 63.3%; Score 38; DB 2; Length 212;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SWEYSVWLSN 10
Db 113 SWTFSTWVHN 122
|||:|||||

RESULT 11

F90331
partial transposase ISCl359 [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 18-Jul-2001

C;Accession: F90331; F90341
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: F90331
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-251 <KUR>
A;Cross-references: GB:AE006641; NID:gl3814951; PIDN:AAK41909.1; GSPDB:GN00155
C;Accession: F90341
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-251 <KU2>
A;Cross-references: GB:AE006641; NID:gl3815045; PIDN:AAK41989.1; GSPDB:GN00155
C;Genetics:
A;Gene: SS01703; SS01796

Query Match 63.3%; Score 38; DB 2; Length 251;
Best Local Similarity 77.8%; Pred. No. 42;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 WEYSVWLSN 10
||:||||
Db 221 WEFSVLSN 229

RESULT 12
D90345
partial transposase ISCI359 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: D90345
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-288 <KUR>
A;Cross-references: GB:AE006641; NID:gl3815079; PIDN:AAK42019.1; GSPDB:GN00155
C;Genetics:
A;Gene: SS01822

Query Match 63.3%; Score 38; DB 2; Length 288;
Best Local Similarity 77.8%; Pred. No. 48;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 WEYSVWLSN 10
||:||||
Db 258 WEFSVLSN 266

RESULT 13
E90334
partial transposase ISCI359 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C;Accession: E90334
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: E90334
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-310 <KUR>
A;Cross-references: GB:AE006641; NID:gl3814978; PIDN:AAK41932.1; GSPDB:GN00155
C;Genetics:
A;Gene: SS01735

Query Match 63.3%; Score 38; DB 2; Length 310;
Best Local Similarity 77.8%; Pred. No. 52;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 WEYSVWLSN 10
||:||||
Db 280 WEFSVLSN 288

RESULT 14
C90497
transposase ISCI359 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C;Accession: C90497
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: C90497
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-332 <KUR>
A;Cross-references: GB:AE006641; NID:gl3816557; PIDN:AAK43234.1; GSPDB:GN00155
C;Genetics:
A;Gene: SS03135

Query Match 63.3%; Score 38; DB 2; Length 332;
Best Local Similarity 77.8%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 WEYSVWLSN 10
||:||||
Db 320 WEFSVLSN 328

RESULT 15
A90279
transposase ISCI359 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C;Accession: A90279
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: A90279
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-350 <KUR>
A;Cross-references: GB:AE006641; NID:gl3814442; PIDN:AAK41488.1; GSPDB:GN00155
C;Genetics:
A;Gene: SS01245

Query Match 63.3%; Score 38; DB 2; Length 350;
Best Local Similarity 77.8%; Pred. No. 59;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 WEYSVWLSN 10
||:||||
Db 320 WEFSVLSN 328

Search completed: May 19, 2004, 15:47:10

Thu May 20 11:48:30 2004

us-10-005-684-5.rpr

Page 5

Job time : 8.29167 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 15:27:51 ; Search time 22.5 seconds
(without alignments)
123.672 Million cell updates/sec

Title: US-10-005-684-5
Perfect score: 60
Sequence: 1 SWEYSVWLSN 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1145568 seqs, 278261457 residues

Total number of hits satisfying chosen parameters: 1145568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 60 | 100.0 | 10 | 14 | US-10-005-684-5 |
| 2 | 56 | 93.3 | 10 | 14 | US-10-005-684-3 |
| 3 | 41 | 68.3 | 812 | 15 | US-10-369-493-9134 |
| 4 | 39 | 65.0 | 82 | 9 | US-09-864-761-41255 |
| 5 | 39 | 65.0 | 305 | 12 | US-10-267-502-389 |
| 6 | 39 | 65.0 | 450 | 14 | US-10-156-761-14443 |
| 7 | 38 | 63.3 | 80 | 9 | US-09-764-860-381 |
| 8 | 38 | 63.3 | 80 | 14 | US-10-074-095-381 |
| 9 | 38 | 63.3 | 80 | 15 | US-10-212-872-381 |
| 10 | 38 | 63.3 | 213 | 15 | US-10-320-797-3029 |
| 11 | 38 | 63.3 | 645 | 14 | US-10-270-333-180 |
| 12 | 38 | 63.3 | 654 | 14 | US-10-246-330-28 |
| 13 | 37 | 61.7 | 71 | 11 | US-09-864-408A-2112 |
| 14 | 37 | 61.7 | 532 | 12 | US-10-425-114-70367 |
| 15 | 37 | 61.7 | 535 | 12 | US-10-425-114-51560 |

| | | | | | | |
|----|----|------|------|----|----------------------|---------------------|
| 16 | 37 | 61.7 | 541 | 12 | US-10-425-114-70602 | Sequence 70602, A |
| 17 | 37 | 61.7 | 544 | 12 | US-10-424-599-185946 | Sequence 185946 |
| 18 | 37 | 61.7 | 1848 | 10 | US-09-839-996-6 | Sequence 6, Appli |
| 19 | 37 | 61.7 | 1848 | 12 | US-10-645-655-6 | Sequence 6, Appli |
| 20 | 37 | 61.7 | 1848 | 14 | US-10-080-505-6 | Sequence 208, Appli |
| 21 | 36 | 60.0 | 43 | 10 | US-09-986-480-208 | Sequence 283362, |
| 22 | 36 | 60.0 | 54 | 12 | US-10-424-599-283362 | Sequence 283362, |
| 23 | 36 | 60.0 | 55 | 12 | US-10-424-599-232805 | Sequence 232805, |
| 24 | 36 | 60.0 | 76 | 12 | US-10-424-599-271499 | Sequence 271499, |
| 25 | 36 | 60.0 | 79 | 12 | US-10-424-599-267546 | Sequence 267546, |
| 26 | 36 | 60.0 | 103 | 10 | US-09-764-891-3652 | Sequence 3652, Ap |
| 27 | 36 | 60.0 | 139 | 14 | US-10-156-761-10961 | Sequence 10961, A |
| 28 | 36 | 60.0 | 161 | 14 | US-10-253-971-2 | Sequence 2, Appli |
| 29 | 36 | 60.0 | 161 | 15 | US-10-439-478-6 | Sequence 6, Appli |
| 30 | 36 | 60.0 | 194 | 14 | US-10-128-714-3266 | Sequence 3266, Ap |
| 31 | 36 | 60.0 | 219 | 15 | US-10-369-493-12382 | Sequence 12382, A |
| 32 | 36 | 60.0 | 250 | 14 | US-10-128-714-8266 | Sequence 8266, Ap |
| 33 | 36 | 60.0 | 263 | 12 | US-10-282-122A-59487 | Sequence 59487, A |
| 34 | 36 | 60.0 | 875 | 12 | US-10-282-122A-47255 | Sequence 47255, A |
| 35 | 36 | 60.0 | 1185 | 12 | US-10-282-122A-46181 | Sequence 46181, A |
| 36 | 36 | 60.0 | 1219 | 15 | US-10-369-493-3212 | Sequence 3212, Ap |
| 37 | 36 | 60.0 | 1267 | 12 | US-10-282-122A-68187 | Sequence 68187, A |
| 38 | 36 | 60.0 | 1541 | 10 | US-09-839-996-3 | Sequence 3, Appli |
| 39 | 36 | 60.0 | 1541 | 12 | US-10-645-655-3 | Sequence 3, Appli |
| 40 | 36 | 60.0 | 1541 | 14 | US-10-080-505-3 | Sequence 3, Appli |
| 41 | 36 | 60.0 | 1545 | 10 | US-09-839-996-4 | Sequence 4, Appli |
| 42 | 36 | 60.0 | 1545 | 12 | US-10-645-655-4 | Sequence 4, Appli |
| 43 | 36 | 60.0 | 1545 | 14 | US-10-080-505-4 | Sequence 4, Appli |
| 44 | 36 | 60.0 | 1694 | 12 | US-10-282-122A-58377 | Sequence 58377, A |
| 45 | 36 | 60.0 | 1702 | 10 | US-09-839-996-5 | Sequence 5, Appli |

ALIGNMENTS

RESULT 1
US-10-005-684-5
; Sequence 5, Application US/10005684
; Publication No. US20030100035A1
; GENERAL INFORMATION:
; APPLICANT: Vojdani, Aristo
; TITLE OF INVENTION: SALIVA IMMUNOASSAY FOR DETECTION OF
; TITLE OF INVENTION: ANTIBODIES FOR AUTOIMMUNE DISEASE
; FILE REFERENCE: IMSC12.005A
; CURRENT APPLICATION NUMBER: US/10/005,684
; CURRENT FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Lupus peptide
US-10-005-684-5

Query Match 100.0%; Score 60; DB 14; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SWEYSVWLSN 10
DB 1 SWEYSVWLSN 10

RESULT 2
US-10-005-684-3
; Sequence 3, Application US/10005684
; Publication No. US20030100035A1
; GENERAL INFORMATION:
; APPLICANT: Vojdani, Aristo
; TITLE OF INVENTION: SALIVA IMMUNOASSAY FOR DETECTION OF
; TITLE OF INVENTION: ANTIBODIES FOR AUTOIMMUNE DISEASE
; FILE REFERENCE: IMSC12.005A
; CURRENT APPLICATION NUMBER: US/10/005,684

```
; CURRENT FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 10
; TYPE: PRT
; ORGANISM: lupus-like autoimmunity inducing peptide
US-10-005-684-3

Query Match
Best Local Similarity 93.3%; Score 56; DB 14; Length 10;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 WEYSVWLSN 10
Db 2 WEYSVWLSN 10

RESULT 3
US-10-369-493-9134
; Sequence 9134, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9134
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Chloroflexus aurantiacus
US-10-369-493-9134

Query Match
Best Local Similarity 68.3%; Score 41; DB 15; Length 812;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 WEYSVW 7
Db 415 WEYSIW 420

RESULT 4
US-09-864-761-41255
; Sequence 41255, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecwica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
```

```
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41255
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL078646.29
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EST HUMAN HIT: AA642505.1, EVALUE 2.30e+00
; OTHER INFORMATION: SWISSPROT HIT: P44044, EVALUE 5.70e+00
US-09-864-761-41255

Query Match
Best Local Similarity 65.0%; Score 39; DB 9; Length 82;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 WEYSVWLSN 10
Db 66 TWBGIWVRN 75

RESULT 5
US-10-267-502-389
; Sequence 389, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: USD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 389
; LENGTH: 305
; TYPE: PRT
```

```
; ORGANISM: Drosophila melanogaster
US-10-267-502-389

Query Match      65.0%; Score 39; DB 12; Length 305;
Best Local Similarity 50.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 SWEYSVWLSN 10
Db      71 SWEFEIWGN 80

RESULT 6
US-10-156-761-14443
; Sequence 14443, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14443
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14443

Query Match      55.0%; Score 39; DB 14; Length 450;
Best Local Similarity 55.8%; Pred. No. 6.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 SWEYSVWLS 9
Db      94 TWYATWTS 102

RESULT 7
US-09-764-860-381
; Sequence 381, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 381
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (49)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-860-381

Query Match      63.3%; Score 38; DB 9; Length 80;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
```

```
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 SWEYSVWLS 9
Db      64 SWEYSVWVS 72

RESULT 8
US-10-074-095-381
; Sequence 381, Application US/10074095
; Publication No. US20030077704A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C1
; CURRENT APPLICATION NUMBER: US/10/074,095
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: 09/764,860
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
```

; PRIOR APPLICATION NUMBER: 60/241,785
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/244,617
 ; PRIOR FILING DATE: 2000-11-01
 ; PRIOR APPLICATION NUMBER: 60/225,268
 ; PRIOR FILING DATE: 2000-08-14
 ; PRIOR APPLICATION NUMBER: 60/236,368
 ; PRIOR FILING DATE: 2000-09-29
 ; PRIOR APPLICATION NUMBER: 60/251,856
 ; PRIOR FILING DATE: 2000-12-08
 ; PRIOR APPLICATION NUMBER: 60/251,868
 ; PRIOR FILING DATE: 2000-12-08
 ; PRIOR APPLICATION NUMBER: 60/229,344
 ; PRIOR FILING DATE: 2000-09-01
 ; PRIOR APPLICATION NUMBER: 60/234,997
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: 60/229,343
 ; PRIOR FILING DATE: 2000-09-01
 ; PRIOR APPLICATION NUMBER: 60/229,345
 ; PRIOR FILING DATE: 2000-09-01
 ; PRIOR APPLICATION NUMBER: 60/229,287
 ; PRIOR FILING DATE: 2000-09-01
 ; PRIOR APPLICATION NUMBER: 60/229,513
 ; PRIOR FILING DATE: 2000-09-05
 ; PRIOR APPLICATION NUMBER: 60/231,413
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 60/229,509
 ; PRIOR FILING DATE: 2000-09-05
 ; PRIOR APPLICATION NUMBER: 60/236,367
 ; PRIOR FILING DATE: 2000-09-29
 ; PRIOR APPLICATION NUMBER: 60/237,039
 ; PRIOR FILING DATE: 2000-10-02
 ; PRIOR APPLICATION NUMBER: 60/237,038
 ; PRIOR FILING DATE: 2000-10-02
 ; PRIOR APPLICATION NUMBER: 60/236,370
 ; PRIOR FILING DATE: 2000-09-29
 ; PRIOR APPLICATION NUMBER: 60/236,802
 ; PRIOR FILING DATE: 2000-10-02
 ; PRIOR APPLICATION NUMBER: 60/237,037
 ; PRIOR FILING DATE: 2000-10-02
 ; PRIOR APPLICATION NUMBER: 60/237,040
 ; PRIOR FILING DATE: 2000-10-02
 ; PRIOR APPLICATION NUMBER: 60/240,960
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/239,935
 ; PRIOR FILING DATE: 2000-10-13
 ; PRIOR APPLICATION NUMBER: 60/239,937
 ; PRIOR FILING DATE: 2000-10-13
 ; PRIOR APPLICATION NUMBER: 60/241,787
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/246,474
 ; PRIOR FILING DATE: 2000-11-08
 ; PRIOR APPLICATION NUMBER: 60/246,532
 ; PRIOR FILING DATE: 2000-11-08
 ; PRIOR APPLICATION NUMBER: 60/249,216
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,210
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/226,681
 ; PRIOR FILING DATE: 2000-08-22
 ; PRIOR APPLICATION NUMBER: 60/225,759
 ; PRIOR FILING DATE: 2000-08-14
 ; PRIOR APPLICATION NUMBER: 60/225,213
 ; PRIOR FILING DATE: 2000-08-14
 ; PRIOR APPLICATION NUMBER: 60/227,182
 ; PRIOR FILING DATE: 2000-08-22
 ; PRIOR APPLICATION NUMBER: 60/225,214
 ; PRIOR FILING DATE: 2000-08-14
 ; PRIOR APPLICATION NUMBER: 60/235,836
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: 60/230,438
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/215,135

; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: 60/225,266
 ; PRIOR FILING DATE: 2000-08-14
 ; PRIOR APPLICATION NUMBER: 60/249,218
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,208
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,213
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,212
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,207
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,245
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,244
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,217
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,211
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,215
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,264
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,214
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/249,297
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: 60/232,400
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/231,242
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 60/232,081
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 60/232,080
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 60/231,414
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 60/231,244
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 60/233,064
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/233,063
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/232,397
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/232,399
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/232,401
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/241,808
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/241,826
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/241,786
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/241,221
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/246,475
 ; PRIOR FILING DATE: 2000-11-08
 ; PRIOR APPLICATION NUMBER: 60/231,243
 ; PRIOR FILING DATE: 2000-09-08

Query Match 63.3%; Score 38; DB 14; Length 80;

Best Local Similarity 66.7%; Pred.No. 1.9e+02; Matches 6; Conservative 1; Mismatches 0; Gaps 0;

QY 1 SWEYSVWLS 9

DB 64 SHVSVWVS 72

```
RESULT 9
US-10-212-872-381
; Sequence 381, Application US/10212872
; Publication No. US20030215893A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008C2
; CURRENT APPLICATION NUMBER: US/10/212,872
; CURRENT FILING DATE: 2002-08-07
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 381
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (49)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-212-872-381

Query Match      63.3%; Score 38; DB 15; Length 80;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 SWEYSWLS 9
Db      64 SWHVSWS 72

RESULT 10
US-10-320-797-3029
; Sequence 3029, Application US/10320797
; Publication No. US20040014955A1
; GENERAL INFORMATION:
; APPLICANT: Eroshkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3029
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Cryptococcus neoformans
US-10-320-797-3029

Query Match      63.3%; Score 38; DB 15; Length 213;
Best Local Similarity 40.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      1 SWEYSWLSN 10
Db      120 TWSFSIWLS 129

RESULT 11
US-10-270-333-180
; Sequence 180, Application US/10270333
; Publication No. US20030032124A1
; GENERAL INFORMATION:
; APPLICANT: Cravchik, Anibal
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF AS INSECTICIDAL TARGETS
; FILE REFERENCE: CL000733CON

; CURRENT APPLICATION NUMBER: US/10/270,333
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/168,677
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/175,691
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/191,638
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 180
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Drosophila
US-10-270-333-180

Query Match      63.3%; Score 38; DB 14; Length 645;
Best Local Similarity 55.6%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 WEYSWLSN 10
Db      403 WNYETWLS 411

RESULT 12
US-10-246-330-28
; Sequence 28, Application US/10246330
; Publication No. US20030166030A1
; GENERAL INFORMATION:
; APPLICANT: O'Toole, George A.
; APPLICANT: Mah, Thien-Fah
; TITLE OF INVENTION: METHODS TO STUDY AND MECHANISMS OF
; TITLE OF INVENTION: BIOFILM-SPECIFIC ANTIBIOTIC RESISTANCE
; FILE REFERENCE: 14537-002001
; CURRENT APPLICATION NUMBER: US/10/246,330
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 60/323,241
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 654
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-10-246-330-28

Query Match      63.3%; Score 38; DB 14; Length 654;
Best Local Similarity 62.5%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 WEYSWLS 9
Db      422 WTYSTWIS 429

RESULT 13
US-09-864-408A-2112
; Sequence 2112, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1el Human Polynucleotides and Polypeptides Enc
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2112
; LENGTH: 71
```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-2112

Query Match 61.7%; Score 37; DB 11; Length 71;
Best Local Similarity 60.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SWEYSVWLSN 10
Db 44 SWQNSQWTSN 53

RESULT 14

US-10-425-114-70367
; Sequence 70367, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70367
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: CC-ZMLE18900031H12_FLI.pep
US-10-425-114-70367

Query Match 61.7%; Score 37; DB 12; Length 532;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SWEYSVWL 8
Db 139 AWDYSAWV 146

RESULT 15

US-10-425-114-51560
; Sequence 51560, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 51560
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700450428_FLI.pep
US-10-425-114-51560

Query Match 61.7%; Score 37; DB 12; Length 535;

Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SWEYSVWL 8
Db 140 AWDYSAWV 147

Search completed: May 19, 2004, 15:41:16
Job time : 22.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using ew model

Run on: May 19, 2004, 15:23:26 ; Search time 8.75 seconds
(without alignments)
59.001 Million cell updates/sec

Title: US-10-005-684-5
Perfect score: 60
Sequence: 1 SWEYSWLSN 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 56 | 93.3 | 10 | 3 | US-08-531-832A-1 |
| 2 | 43 | 71.7 | 11 | 2 | US-08-665-202-13 |
| 3 | 43 | 71.7 | 11 | 4 | US-09-315-574-13 |
| 4 | 39 | 65.0 | 288 | 4 | US-09-252-991A-17252 |
| 5 | 39 | 65.0 | 405 | 4 | US-08-311-731A-12 |
| 6 | 38 | 63.3 | 545 | 4 | US-09-252-991A-25304 |
| 7 | 38 | 63.3 | 641 | 4 | US-09-489-039A-12721 |
| 8 | 37 | 61.7 | 11 | 2 | US-08-665-202-10 |
| 9 | 37 | 61.7 | 11 | 4 | US-09-315-574-10 |
| 10 | 37 | 61.7 | 1848 | 4 | US-08-839-996-6 |
| 11 | 37 | 61.7 | 1848 | 4 | US-08-839-996-6 |
| 12 | 37 | 61.7 | 1848 | 4 | US-10-080-505-6 |
| 13 | 37 | 61.7 | 1848 | 5 | PCT-US95-10661A-6 |
| 14 | 36 | 60.0 | 214 | 4 | US-09-107-532A-6763 |
| 15 | 36 | 60.0 | 289 | 4 | US-09-489-039A-12390 |
| 16 | 36 | 60.0 | 358 | 4 | US-09-328-352-6447 |
| 17 | 36 | 60.0 | 1541 | 3 | US-08-296-791-3 |
| 18 | 36 | 60.0 | 1541 | 4 | US-08-839-996-3 |
| 19 | 36 | 60.0 | 1541 | 4 | US-10-080-505-3 |
| 20 | 36 | 60.0 | 1541 | 5 | PCT-US95-10661A-3 |
| 21 | 36 | 60.0 | 1545 | 3 | US-08-296-791-4 |
| 22 | 36 | 60.0 | 1545 | 4 | US-09-839-996-4 |
| 23 | 36 | 60.0 | 1545 | 4 | US-10-080-505-4 |
| 24 | 36 | 60.0 | 1545 | 5 | PCT-US95-10661A-4 |
| 25 | 36 | 60.0 | 1702 | 3 | US-08-296-791-5 |
| 26 | 36 | 60.0 | 1702 | 4 | US-08-839-996-5 |
| 27 | 36 | 60.0 | 1702 | 4 | US-10-080-505-5 |

| | | | | | |
|----|----|------|------|---|-------------------|
| 28 | 36 | 60.0 | 1702 | 5 | PCT-US95-10661A-5 |
| 29 | 35 | 58.3 | 11 | 2 | US-08-665-202-20 |
| 30 | 35 | 58.3 | 11 | 4 | US-09-315-574-20 |
| 31 | 35 | 58.3 | 102 | 3 | US-08-894-173-75 |
| 32 | 35 | 58.3 | 102 | 3 | US-09-398-193-75 |
| 33 | 35 | 58.3 | 103 | 3 | US-08-894-173-77 |
| 34 | 35 | 58.3 | 103 | 3 | US-08-894-173-78 |
| 35 | 35 | 58.3 | 103 | 3 | US-09-398-193-77 |
| 36 | 35 | 58.3 | 103 | 3 | US-09-398-193-78 |
| 37 | 35 | 58.3 | 104 | 3 | US-08-894-173-70 |
| 38 | 35 | 58.3 | 104 | 3 | US-08-894-173-71 |
| 39 | 35 | 58.3 | 104 | 3 | US-09-398-193-70 |
| 40 | 35 | 58.3 | 104 | 3 | US-09-398-193-71 |
| 41 | 35 | 58.3 | 109 | 3 | US-08-894-173-95 |
| 42 | 35 | 58.3 | 109 | 3 | US-08-894-173-96 |
| 43 | 35 | 58.3 | 109 | 3 | US-09-398-193-95 |
| 44 | 35 | 58.3 | 109 | 3 | US-09-398-193-96 |
| 45 | 35 | 58.3 | 110 | 3 | US-08-894-173-87 |

ALIGNMENTS

RESULT 1
US-08-531-832A-1
; Sequence 1, Application US/08531832A
; Patent No. 6001964
; GENERAL INFORMATION:
; APPLICANT: BRUCE D. GAYNOR, BETTY A. DIAMOND,
; APPLICANT: MATTHEW D. SCHARFF AND PHILIPPE VALADON
; TITLE OF INVENTION: PEPTIDES FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF SYSTEMIC LUPUS
; TITLE OF INVENTION: ERYTHEMATOSUS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN
; STREET: 90 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,832A
; FILING DATE: SEPTEMBER 20, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG J. ARNOLD
; REGISTRATION NUMBER: 34,287
; REFERENCE/DOCKET NUMBER: 96700/342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082
; TELEX: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; DESCRIPTION: YES
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; US-08-531-832A-1

Query Match 93.3% Score 56; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0064; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0

```
QY      2 WEYSVWLSN 10
Db      2 WEYSVWLSN 10
|||||
|||||

RESULT 2
US-08-665-202-13
; Sequence 13, Application US/08665202
; Patent No. 5977322
; GENERAL INFORMATION:
; APPLICANT: Marks, James D.
; APPLICANT: Schier, Robert
; TITLE OF INVENTION: No. 5977322el High Affinity Human Antibodies to
; TITLE OF INVENTION: Tumor Antigens
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,202
; FILING DATE: 13-JUN-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,238
; FILING DATE: 14-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,250
; FILING DATE: 15-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/665,202
; FILING DATE: 13-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 02307E-061410
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-665-202-13

Query Match      71.7%; Score 43; DB 2; Length 11;
Best Local Similarity 85.7%; Pred. No. 0.7;
Matches      6; Conservative      1; Mismatches      0; Indels      0; Gaps      0;

QY      1 SWEYSVM 7
Db      2 SWEYSVM 8
|||||

RESULT 3
US-09-315-574-13
; Sequence 13, Application US/09315574
; Patent No. 6512097
; GENERAL INFORMATION:
; APPLICANT: Marks, James D.
; APPLICANT: Schier, Robert
; TITLE OF INVENTION: No. 6512097el High Affinity Human Antibodies to
; TITLE OF INVENTION: Tumor Antigens
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Majestic, Parsons, Siebert & Haue P.C.
; STREET: Four Embarcadero Center, Suite 1100
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4106
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,574
; FILING DATE: 20-MAY-99
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,238
; FILING DATE: 14-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,250
; FILING DATE: 15-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/665,202
; FILING DATE: 13-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 02307E-061411
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-315-574-13

Query Match      71.7%; Score 43; DB 4; Length 11;
Best Local Similarity 85.7%; Pred. No. 0.7;
Matches      6; Conservative      1; Mismatches      0; Indels      0; Gaps      0;

QY      1 SWEYSVM 7
Db      2 SWEYSVM 8
|||||

RESULT 4
US-09-252-991A-17252
; Sequence 17252, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17252
; LENGTH: 288
; TYPE: PRT
; ORGANISM: pseudomonas aeruginosa
US-09-252-991A-17252

Query Match      65.0%; Score 39; DB 4; Length 288;
Best Local Similarity 50.0%; Pred. No. 77;
```

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SWEYSVWLSN 10
Db 242 SWFFSAMWN 251

RESULT 5

US-08-311-731A-12
; Sequence 12, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/311.731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 406 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM TUBERCULOSIS

US-08-311-731A-12

Query Match 65.0%; Score 39; DB 4; Length 406;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SWEYSVWLS 9
Db 352 SWNYTTWLN 360

RESULT 6

US-09-252-991A-25304
; Sequence 25304, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25304
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25304

Query Match 63.3%; Score 38; DB 4; Length 545;
Best Local Similarity 55.8%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 WEYSVWLSN 10
Db 46 WDLSLWLN 54

RESULT 7

US-09-489-039A-12721
; Sequence 12721, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12721
; LENGTH: 641
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-12721

Query Match 63.3%; Score 38; DB 4; Length 641;
Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 SWEYSVWLSN 10
Db 191 NWKFDWLEN 200

RESULT 8

US-08-665-202-10
; Sequence 10, Application US/08665202
; Patent No. 5977322
; GENERAL INFORMATION:
; APPLICANT: Marks, James D.
; APPLICANT: Schier, Robert
; TITLE OF INVENTION: No. 5977322el High Affinity Human Antibodies to
; TITLE OF INVENTION: Tumor Antigens
; NUMBER OF SEQUENCES: 141
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/665,202
; FILING DATE: 13-JUN-1996

CLASSIFICATION: 424
PRIOR APPLICATION DATA: US 60/000,238
FILING DATE: 14-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,250
FILING DATE: 15-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 02307B-061410
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-665-202-10

Query Match 61.7%; Score 37; DB 2; Length 11;
Best Local Similarity 57.1%; Pred. No. 5.8;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SWEYSVM 7
Db 2 AWDYSLM 8

RESULT 9
US-09-315-574-10
Sequence 10, Application US/09315574
Patent No. 6512097
GENERAL INFORMATION:
APPLICANT: Marks, James D.
APPLICANT: Schier, Robert
TITLE OF INVENTION: No. 6512097el High Affinity Human Antibodies to
Tumor Antigens
NUMBER OF SEQUENCES: 141
CORRESPONDENCE ADDRESS:
ADDRESSEE: Majestic, Parsons, Siebert & Hsue P.C.
STREET: Four Embarcadero Center, Suite 1100
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4106
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/315,574
FILING DATE: 20-MAY-99
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,238
FILING DATE: 14-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,250
FILING DATE: 15-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/665,202
FILING DATE: 13-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 02307B-061411
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-315-574-10

Query Match 61.7%; Score 37; DB 4; Length 11;
Best Local Similarity 57.1%; Pred. No. 5.8;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 SWEYSVM 7
Db 2 AWDYSLM 8

RESULT 10
US-08-296-791-6
Sequence 6, Application US/08296791
Patent No. 6245337
GENERAL INFORMATION:
APPLICANT: St. Geme III, Joseph W.
APPLICANT: Falkow, Stanley
TITLE OF INVENTION: Haemophilus Adherence and Penetration
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohnbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/296,791
FILING DATE: 25-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277239
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1848 amino acids
TYPE: amino acid
TOPOLOGY: unknown
US-08-296-791-6

Query Match 61.7%; Score 37; DB 3; Length 1848;
Best Local Similarity 62.5%; Pred. No. 1e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 EYSVWSN 10
Db 1600 QYNWWSN 1607

RESULT 11
US-09-839-996-6
Sequence 6, Application US/09839996
Patent No. 6642371

```

; GENERAL INFORMATION:
; APPLICANT: St. Geme III, Joseph W.
; TITLE OF INVENTION: Haemophilus Adherence and Penetration
;   Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/839,996
; FILING DATE: 20-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 398-3249
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1848 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-839-996-6

Query Match 61.7%; Score 37; DB 4; Length 1848;
Best Local Similarity 62.5%; Pred. No. 1e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 EYSVWLSN 10
Db 1600 QYNVWISN 1607

RESULT 12
US-10-080-505-6
; Sequence 6, Application US/10080505
; Patent No. 6676948
; GENERAL INFORMATION:
; APPLICANT: St. Geme, Joseph W.
; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS
; FILE REFERENCE: A-59941-1/RFT/DCF/DHR
; CURRENT APPLICATION NUMBER: US/10/080,505
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: US 08/296,791
; PRIOR FILING DATE: 1994-10-25
; PRIOR APPLICATION NUMBER: US 09/839,996
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1848
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-080-505-6

Query Match 61.7%; Score 37; DB 4; Length 1848;

; GENERAL INFORMATION:
; APPLICANT: St. Geme III, Joseph W.
; TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/839,996
; FILING DATE: 20-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 398-3249
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1848 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-839-996-6

Query Match 61.7%; Score 37; DB 5; Length 1848;
Best Local Similarity 62.5%; Pred. No. 1e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 EYSVWLSN 10
Db 1600 QYNVWISN 1607

RESULT 13
PCT-US95-10661A-6
; Sequence 6, Application PC/TUS9510661A
; GENERAL INFORMATION:
; APPLICANT: Washington University, et al.
; TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10661A
; FILING DATE: 16-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: FP-59941/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1848 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; PCT-US95-10661A-6

Query Match 61.7%; Score 37; DB 5; Length 1848;
Best Local Similarity 62.5%; Pred. No. 1e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 EYSVWLSN 10
Db 1600 QYNVWISN 1607

RESULT 14
US-09-107-532A-6763
; Sequence 6763, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; PCT-US95-10661A-6
```

COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6763:
SEQUENCE CHARACTERISTICS:
LENGTH: 214 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...214
SEQUENCE DESCRIPTION: SEQ ID NO: 6763:
US-09-107-532A-6763

Query Match 60.0%; Score 36; DB 4; Length 214;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SWEYSVWLSN 10
DB 116 SWNLSTWIHN 125

RESULT 15
US-09-489-039A-12390
; Sequence 12390, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12390
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12390

Query Match 60.0%; Score 36; DB 4; Length 289;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SWEYSVW 7
DB 208 SWAYGVW 214

Search completed: May 19, 2004, 15:29:40
Job time : 9.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 15:28:52 ; Search time 29.7917 Seconds
(without alignments)
94.841 Million cell updates/sec

Title: US-10-005-684-5
Perfect score: 60
Sequence: 1 SWEYSVWLSN 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 60 | 100.0 | 10 | 7 ADB81378 | ADB81378 Lupus pep |
| 2 | 56 | 93.3 | 10 | 3 AAY51201 | AAY51201 Human ant |
| 3 | 43 | 71.7 | 11 | 2 AAW08517 | Aaw08517 C6 human |
| 4 | 41 | 68.3 | 228 | 4 ABG06373 | ABg06373 Novel hum |
| 5 | 41 | 68.3 | 2570 | 4 ABG08375 | ABg08375 Novel hum |
| 6 | 40 | 66.7 | 539 | 3 AAY74632 | Aay74632 Neisseria |
| 7 | 39 | 65.0 | 82 | 4 AAM20846 | Aam20846 Peptide # |
| 8 | 39 | 65.0 | 82 | 4 ABB42579 | Abb42579 Peptide # |
| 9 | 39 | 65.0 | 82 | 4 AAM36395 | Aam36395 Peptide # |
| 10 | 39 | 65.0 | 82 | 4 ABB25957 | Abb25957 Protein # |
| 11 | 39 | 65.0 | 82 | 4 AAM76283 | Aam76283 Human bon |
| 12 | 39 | 65.0 | 82 | 4 AAM63469 | Aam63469 Human bra |
| 13 | 39 | 65.0 | 82 | 4 ABG57995 | ABg57995 Human liv |
| 14 | 39 | 65.0 | 82 | 5 ABG45592 | ABg45592 Human pep |
| 15 | 39 | 65.0 | 305 | 4 ABB71625 | Abb71625 Drosophil |
| 16 | 39 | 65.0 | 406 | 7 ADB74263 | ADB74263 Mycobacte |
| 17 | 39 | 65.0 | 1196 | 2 AAW04326 | Aaw04326 Rat petri |
| 18 | 38 | 63.3 | 62 | 4 AAU53331 | AAu53331 Propionib |
| 19 | 38 | 63.3 | 62 | 6 ABM49850 | ABm49850 Propionib |
| 20 | 38 | 63.3 | 80 | 4 AAU17763 | AAu17763 Novel hum |
| 21 | 38 | 63.3 | 85 | 4 ABG15794 | ABg15794 Novel hum |
| 22 | 38 | 63.3 | 95 | 4 AAO06048 | AAo06048 Human pol |
| 23 | 38 | 63.3 | 212 | 5 ABB48467 | ABb48467 Listeria |
| 24 | 38 | 63.3 | 213 | 7 ADB69985 | ADB69985 C. neofo |
| 25 | 38 | 63.3 | 413 | 4 ABB66175 | ABb66175 Drosophil |

ALIGNMENTS

RESULT 1

ADB81378
ID ADB81378 standard; peptide; 10 AA.

XX AC ADB81378;

XX DT 04-DEC-2003 (first entry)

XX DE Lupus peptide (10-mer) used in a test for antibodies against lupus.

XX KW immunoassay; autoimmune disease; autoantigen; ELISA;
KW enzyme-linked immunosorbant assay; saliva IgA; lupus; arthritis;
KW platelet glycoprotein; immune complex.

XX OS Synthetic.

XX FN US2003100035-A1.

XX PD 29-MAY-2003.

XX PF 08-NOV-2001; 2001US-00005684.

XX PR 08-NOV-2001; 2001US-00005684.

XX PA (VOJD/) VOJDANI A.

XX PI Vojdani A;

XX XX WPI; 2003-606630/57.

XX DR Saliva immunoassay for detection of antibodies for autoimmune disease,
XX PT e.g. lupus, in patient, by determining level of antibodies against
XX PT autoantigen, and comparing level of determined antibodies with normal
XX PT levels of antibodies.

XX XX Example 8; Page 7; 16pp; English.

XX CC This invention relates to a novel saliva immunoassay for diagnosing an
XX CC autoimmune disease in a patient. Specifically, the method determines the
XX CC level of antibodies that are present against the autoantigens for a
XX CC particular autoimmune disease and compares these results with normal
XX CC levels to determine the likelihood or severity of such a disease. The
XX CC test comprises a highly sensitive and accurate ELISA (enzyme-linked
XX CC immunosorbant assay) that measures saliva IgA specific antibody titres
XX CC against the purified antigens or a corresponding recombinant antigen or
XX CC synthetic peptide of an autoantigen, where the autoantigens are lupus
XX CC peptides, arthritis peptides, platelet glycoprotein or immune complexes.
XX CC As such, this single test can be used to accurately detect antibodies for

Abb70763 Drosophil
Aau38982 Drosophil
Adc35928 Drosophil
Abp59944 A tumefac
Aae38201 Fruit fly
ABg24833 Novel hum
ABg04046 Novel hum
Aaw08514 C6 human
Aar83353 Kbm1-bind
ABp32083 Human ORF
Aab68933 Neisseria
Aag47140 Arabidops
Aag14083 Arabidops
Aag47139 Arabidops
Aag14082 Arabidops
Aag14081 Arabidops
Aag47138 Arabidops
Aca09347 Haemophil
Aab56114 Human sec
Aab25694 Human sec

26 38 63.3 645 4 ABB70763
27 38 63.3 645 4 AAU38982
28 38 63.3 645 7 ADC35928
29 38 63.3 654 6 ABP59944
30 38 63.3 755 7 AAE38201
31 38 63.3 816 4 ABG24833
32 38 63.3 858 4 ABG04046
33 37 61.7 11 2 AAW08514
34 37 61.7 12 2 AAR83353
35 37 61.7 71 5 ABP32083
36 37 61.7 115 4 AAB68933
37 37 61.7 229 3 AAG47140
38 37 61.7 229 3 AAG14083
39 37 61.7 269 3 AAG47139
40 37 61.7 269 3 AAG14082
41 37 61.7 318 3 AAG14081
42 37 61.7 318 3 AAG47138
43 37 61.7 1848 6 ADA09347
44 36 60.0 43 3 AAB56114
45 36 60.0 44 3 AAB25694

CC diseases such as lupus or arthritis. Furthermore, it is useful to
 CC indicate ongoing pathology or to predict an early pathogenic reaction for
 CC autoimmune disease. This peptide sequence is the lupus peptide (10-mer)
 CC used in a test for lupus antibodies, in an exemplification of the
 CC invention.

XX SQ Sequence 10 AA;

Query Match 100.0%; Score 60; DB 7; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.0033;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SWEYSVWLSN 10
 |||||
 Db 1 SWEYSVWLSN 10

RESULT 2

AAW08517
 ID AAY51201 standard; peptide; 10 AA.

XX AC AAY51201;

XX DT 04-APR-2000 (first entry)

XX DE Human anti-dsDNA monoclonal antibody R4A specific epitope.

XX KW Monoclonal antibody; anti-dsDNA; R4A; human; pathogenesis;
 XX KW systemic lupus erythematosus; treatment; diagnosis.

XX OS Homo sapiens.

XX FN US6001964-A.

XX PD 14-DEC-1999.

XX PF 20-SEP-1995; 95US-00531832.

XX PR 20-SEP-1995; 95US-00531832.

XX FA (YESH) UNIV YESHIVA EINSTEIN COLLEGE.

XX PI Gaynor BD, Valadon P, Diamond BA, Scharff MD;

XX DR WPT; 2000-104899/09.

XX PT Peptide specific for an antibody, useful for treatment and diagnosis of
 XX PT systemic lupus erythematosus.

XX PS Disclosure; Col 13-14; 13pp; English.

XX CC This invention describes a novel purified peptide (I) which binds to an
 CC anti-double-stranded DNA antibody. (I) is capable of neutralizing
 CC antibodies important in the pathogenesis of systemic lupus erythematosus.
 CC (I) is useful for the treatment and diagnosis of systemic lupus
 CC erythematosus. This sequence represents a human anti-dsDNA monoclonal
 CC antibody R4A specific epitope which is used in the method of the
 CC invention

XX SQ Sequence 10 AA;

Query Match 93.3%; Score 56; DB 3; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.014;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 WEYSVWLSN 10
 |||||
 Db 2 WEYSVWLSN 10

RESULT 3

AAW08517
 ID AAW08517 standard; protein; 11 AA.

XX AC AAW08517;
 XX DT 23-SEP-1997 (first entry)
 XX DE C6 human sFv antibody C6.5 variable light chain CDR3 mutant C6ML3-4.

XX KW Tumour; immune response; cytotoxin; carcinoma; breast cancer.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO9700271-A1.

XX PD 03-JAN-1997.

XX PF 13-JUN-1996; 96WO-US010287.

XX PR 14-JUN-1995; 95US-0000238P.

XX PR 15-JUN-1995; 95US-0000250P.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Marks JD, Schier R;

XX DR WPI; 1997-077488/07.

XX PT New C6 human antibody binding specifically to c-erbB-2 - useful for
 XX PT treatment and diagnosis of tumours, with reduced risk of generating
 XX PT immune response.

XX PS Claim 10; Page 24; 117pp; English.

XX CC The present sequence represents a variable light chain CDR3 mutant C6ML3-
 XX CC 4 modified from the derived light chain variable region of the C6 human
 XX CC antibody C6.5. The mutant C6ML3-4 specifically binds to c-erbB-2. A
 XX CC chimeric molecule that binds specifically to tumour cells carrying c-erbB
 XX CC -2 consists of an effector compound attached to a C6 human antibody. If
 XX CC the effector compound is a cytotoxin the chimeric molecule can be used to
 XX CC inhibit growth of c-erbB-2 positive tumours (especially breast and other
 XX CC carcinomas). If the effector compound is a label the chimeric molecule
 XX CC can be used to detect such cells, including in vivo localisation. The
 XX CC antibody can also be used for diagnosis/localisation, in vivo or in
 XX CC vitro, especially by immunoassay. The nucleic acid encoding the antibody,
 XX CC and a nucleic acid encoding a single chain polypeptide with the binding
 XX CC specificity of the antibody and comprising the binding portions of
 XX CC variable regions of light and heavy chains of the antibody, joined by a
 XX CC linker, can be used to produce recombinant proteins by standard methods.
 XX CC Unlike known anti-c-erbB-2 antibodies, C6 antibodies are fully human, so
 XX CC should elicit little, if any, immunogenic response

XX SQ Sequence 11 AA;

Query Match 71.7%; Score 43; DB 2; Length 11;
 Best Local Similarity 85.7%; Pred. No. 1.8;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SWEYSVW 7
 |||||
 Db 2 SWEYSLW 8

RESULT 4

ABG06373
 ID ABG06373 standard; protein; 228 AA.

XX AC ABG06373;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #6364.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS70560.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX Claim 20; SEQ ID NO 36732; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences. ABG0010-ABG30377 Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 228 AA;

Query Match 68.3%; Score 41; DB 4; Length 228;

Best Local Similarity 70.0%; Pred. No. 1.1e+02;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SWEYSVWLSN 10

Db 207 SWEEDVWLVN 216

RESULT 5

ABG06375

ID ABG06375 standard; protein; 2570 AA.

XX AC ABG06375;

XX 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #6366.

XX Human; chromosome mapping; gene mapping; gene therapy; forensics;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

PN WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS70562.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX Claim 20; SEQ ID NO 36734; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences. ABG0010-ABG30377 Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2570 AA;

Query Match 68.3%; Score 41; DB 4; Length 2570;

Best Local Similarity 70.0%; Pred. No. 1.6e+03;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SWEYSVWLSN 10

Db 291 SWEEDVWLVN 300

RESULT 6

AAAY74632

ID AAAY74632 standard; protein; 539 AA.

XX AC AAAY74632;

XX 21-MAR-2000 (first entry)

XX Neisseria meningitidis ORF 212 protein sequence SEQ ID NO:738.

XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.

OS Neisseria meningitidis.

```

XX PN WO9957280-A2.
XX PD
XX PF 11-NOV-1999.
XX PR
XX PP 30-APR-1999; 99WO-US009346.
XX PT
XX PU 01-MAY-1998; 98US-0083758P.
XX PV 31-JUL-1998; 98US-0094865P.
XX PW 02-SEP-1998; 98US-0098994P.
XX PY 02-SEP-1998; 98US-0099062P.
XX PZ 09-OCT-1998; 98US-0103749P.
XX Q1 09-OCT-1998; 98US-0103794P.
XX Q2 09-OCT-1998; 98US-0103796P.
XX Q3 25-FEB-1999; 99US-0121528P.
XX Q4
XX Q5 (CHIR ) CHIRON CORP.
XX Q6 (GENO-) INST GENOMIC RES.
XX Q7
XX Q8 Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
XX Q9 PI Petersen J, Pizzi M, Rappuoli R, Ratti G, Scarlato E, Scarselli M;
XX Q10 PI Tetzelin H, Venter JC;
XX Q11
XX Q12 WPI; 2000-062150/05.
XX Q13 DR N-PSDB; AAZ53394.
XX Q14
XX Q15 Novel Neisserial polypeptides predicted to be useful antigens for
XX Q16 PT vaccines and diagnostics.
XX Q17
XX Q18 Claim 2; Page 490; 1453pp; English.
XX Q19
XX Q20 AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941
XX Q21 CC represent novel Neisseria meningitis and N. gonorrhoea polynucleotides
XX Q22 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
XX Q23 CC PCR primers used in the exemplification of the present invention. The
XX Q24 CC polypeptides, the polynucleotides, antibodies and compositions of the
XX Q25 CC invention can be used as vaccines, as diagnostic reagents, and as
XX Q26 CC immunogenic compositions. The polypeptides can be used in the manufacture
XX Q27 CC of medicaments for treating or preventing infection due to Neisserial
XX Q28 CC bacteria (e.g. meningitis and septicemia), to detect the presence of
XX Q29 CC Neisseria bacteria, or to raise antibodies. They may also be used to
XX Q30 CC screen for agonists or antagonists, which may themselves have use as
XX Q31 CC antibacterial agents. The polynucleotides of the invention may also be
XX Q32 CC used in gene therapy protocols
XX Q33
XX Q34 SQ Sequence 539 AA;
XX Q35
XX Q36 Query Match 66.7%; Score 40; DB 3; Length 539;
XX Q37 ID Best Local Similarity 62.5%; Pred. No. 4e+02;
XX Q38 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX Q39
XX Q40 QY 2 WEYSVWLS 9
XX Q41 DB 271 WOYGVWLN 278
XX Q42
XX Q43 RESULT 7
XX Q44 AAAM20846
XX Q45 ID AAAM20846 standard; protein; 82 AA.
XX Q46
XX Q47 AC AAAM20846;
XX Q48
XX Q49 DT 12-OCT-2001 (first entry)
XX Q50
XX Q51 PE Peptide #7280 encoded by probe for measuring cervical gene expression.
XX Q52
XX Q53 KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX Q54 KW cervical cancer.
XX Q55
XX Q56 OS Homo sapiens.
XX Q57
XX Q58 PN WO200157278-A2.
XX Q59

```

```

PD 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000670.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human cervical epithelial cells.
XX Claim 27; SEQ ID NO 25672; 487pp; English.
XX The present invention relates to human single exon nucleic acid probes
XX (SENP; see AAI10068-AAI28459). The present sequence is a peptide encoded
XX by one such probe. The SENPs are derived from human HeLa cells. The SENPs
XX can be used to produce a single exon microarray, which can be used for
XX measuring human gene expression in a sample derived from human cervical
XX epithelial cells. By measuring gene expression, the probes are therefore
XX useful in grading and/or staging of diseases of the cervix, notably
XX cervical cancer. Note: The sequence data for this patent did not form
XX part of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 82 AA;
XX Query Match 65.0%; Score 39; DB 4; Length 82;
XX Best Local Similarity 40.0%; Pred. No. 70;
XX Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
XX QY 1 SWEYSVWLSN 10
XX DB 66 TWEHGIWVRN 75
XX
XX RESULT 8
XX ABB42579
XX ID ABB42579 standard; peptide; 82 AA.
XX AC ABB42579;
XX DT 04-FEB-2002 (first entry)
XX DE Peptide #1085 encoded by human foetal liver single exon probe.
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX OS Homo sapiens.
XX PN WO200157277-A2.
XX PD 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000669.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX

```

PA (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human fetal liver.
PT Claim 27; SEQ ID NO 35214; 639pp + Sequence Listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for measuring
XX human gene expression in a sample derived from human foetal liver. The
XX single exon nucleic acid probes may be used for predicting, measuring and
XX displaying gene expression in samples derived from human fetal liver. The
XX present sequence is a peptide encoded by a single exon nucleic acid probe
XX of the invention. Note: The sequence data for this patent did not form
XX part of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 82 AA;
SQ
Query Match 65.0%; Score 39; DB 4; Length 82;
Best Local Similarity 40.0%; Pred. No. 70;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Qy 1 SWEYSVWLSN 10
Db 66 TWEHGIWVRN 75
RESULT 9
ID AAM36395 standard; protein; 82 AA.
XX
AC AAM36395;
XX
XX 17-OCT-2001 (first entry)
XX
XX Peptide #10432 encoded by probe for measuring placental gene expression.
DE Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
KW
XX Homo sapiens.
OS
XX WO200157272-A2.
PN
XX
PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207458P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488897/53.
XX
XX Human Genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
PT
XX Claim 27; SEQ ID NO 36664; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP:
XX see AA13115-AA157546). The present sequence is a peptide encoded by one

CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
SQ Sequence 82 AA;
Query Match 65.0%; Score 39; DB 4; Length 82;
Best Local Similarity 40.0%; Pred. No. 70;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Qy 1 SWEYSVWLSN 10
Db 66 TWEHGIWVRN 75
RESULT 10
ID ABB25957 standard; protein; 82 AA.
XX
AC ABB25957;
XX
XX 23-JAN-2002 (first entry)
XX
XX Protein #7956 encoded by probe for measuring heart cell gene expression.
DE Human; Gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
XX Homo sapiens.
OS
XX WO200157274-A2.
PN
XX
PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000666.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207456P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts.
PT
XX Claim 15; SEQ ID NO 27727; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
XX ABA21535-ABA41305). The present sequence is a protein encoded by one such
XX probe. The probes may be used for predicting, measuring and displaying
XX gene expression in samples derived from the human heart via microarrays.
XX By measuring gene expression, the probes are useful for predicting,
XX diagnosing, grading, staging, monitoring and prognosing diseases of the
XX human heart and vascular system e.g. cardiovascular disease,
XX hypertension, cardiac arrhythmias and congenital heart disease. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 82 AA;
SQ
Query Match 65.0%; Score 39; DB 4; Length 82;

PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
PT
XX
XX Claim 27; SEQ ID NO 36643; 659pp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: the sequence
CC information for this patent does not appear in the printed specification
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 82 AA;
XX
XX Query Match 65.0%; Score 39; DB 4; Length 82;
XX Best Local Similarity 40.0%; Pred. No. 70;
XX Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 SWEYSVWLSN 10
XX :||: :||: |
XX 66 TWEHGIWVRN 75
XX
XX RESULT 14
XX ABG45592
XX ID ABG45592 standard; peptide; 82 AA.
XX AC ABG45592;
XX DT 19-AUG-2002 (first entry)
XX
XX Human peptide encoded by genome-derived single exon probe SEQ ID 35257.
XX
XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.
XX
XX Homo sapiens.
XX OS
XX WO200186003-A2.
XX PN
XX 15-NOV-2001.
XX PD
XX 30-JAN-2001; 2001WO-US000665.
XX PF
XX

PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
PT
XX
XX Claim 27; SEQ ID NO 35257; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression to a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberosus sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a peptide/protein encoded by a single exon probe of
CC the invention. Note: the sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 82 AA;
XX
XX Query Match 65.0%; Score 39; DB 5; Length 82;
XX Best Local Similarity 40.0%; Pred. No. 70;
XX Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 SWEYSVWLSN 10
XX :||: :||: |
XX 66 TWEHGIWVRN 75
XX
XX RESULT 15
XX ABB71625
XX ID ABB71625 standard; protein; 305 AA.
XX

AC ABB71625;
 XX 26-MAR-2002 (first entry)
 DT
 DE Drosophila melanogaster polypeptide SEQ ID NO 41667.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEXE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL15728.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 41667; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA
 CC sequences (ABLI1840-ABLI16175) and the encoded proteins (ABBS7737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 305 AA;
 Query Match 65.0%; Score 39; DB 4; Length 305;
 Best Local Similarity 50.0%; Pred. NO. 3e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 SWEYSVWLSN 10
 Db 71 SWEFEIWGQN 80
 Search completed: May 19, 2004, 15:43:54
 Job time : 31.7917 secs